arXiv:math/0101225v1 [math.FA] 26 Jan 2001

BADLY APPROXIMABLE MATRIX FUNCTIONS AND CANONICAL FACTORIZATIONS

R.B. ALEXEEV AND V.V. PELLER

Abstract. We continue studying the problem of analytic approximation of matrix functions. We introduce the notion of a partial canonical factorization of a badly approximable matrix function Φ and the notion of a canonical factorization of a very badly approximable matrix function Φ. Such factorizations are defined in terms of so-called balanced unitary-valued functions which have many remarkable properties. Unlike the case of thematic factorizations studied earlier in [PY1], [PY2], [PT], [AP1], the factors in canonical factorizations (as well as partial canonical factorizations) are uniquely determined by the matrix function Φ up to constant unitary factors. We study many properties of canonical factorizations. In particular we show that under certain natural assumptions on a function space X the condition $\Phi \in X$ implies that all factors in a canonical factorization of Φ belong to the same space X. In the last section we characterize the very badly approximable unitary-valued functions U that satisfy the condition $\|H_U\|_{\mathrm{e}} < 1$.

1. Introduction

The problem of uniform approximation by bounded analytic functions has been studied for a long time. It was proved in [Kh] that for a continuous function φ on the unit circle $\mathbb{T}$ there exists a unique best approximation f by bounded analytic functions and the error function $\varphi - f$ has constant modulus almost everywhere on $\mathbb{T}$. Since that time this approximation problem has been studied by many authors. An important step in developing the theory of approximation by analytic functions was the Nehari theorem [Ne] according to which the distance from an L^∞ function φ to the space H^∞ of bounded analytic functions is equal to the norm of the *Hankel operator* $H_\varphi : H^2 \to H^2_- \stackrel{\text{def}}{=} L^2 \ominus H^2$ defined by

$$H_\varphi f = \mathbb{P}_- \varphi f,$$

where $\mathbb{P}_-$ is the orthogonal projection onto H^2_-. Hankel operators were used essentially for further development of this theory in [AAK1-3] and [PK].

Later it turned out that the approximation problem in question plays a crucial role in so-called H^∞ control theory. Moreover, for the needs of control theory

The second author is partially supported by NSF grant DMS 9970561.

engineers have to consider matrix-valued functions. We refer the reader to [F] for an introduction in H^∞ control.

In this paper we continue the study of best analytic approximation of matrix-valued functions. We consider the space $L^\infty(\mathbb{M}_{m,n})$ of essentially bounded functions which take values in the space $\mathbb{M}_{m,n}$ of $m\times n$ matrices and endow it with the norm

$$\|\Phi\|_{L^\infty} = \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} \|\Phi(\zeta)\|_{\mathbb{M}_{m,n}},$$

(the space $\mathbb{M}_{m,n}$ is endowed with the operator norm in the space of operators from $\mathbb{C}^n$ to $\mathbb{C}^m$) and we study approximations of functions $\Phi\in L^\infty(\mathbb{M}_{m,n})$ by functions in the subspace $H^\infty(\mathbb{M}_{m,n})$ of $L^\infty(\mathbb{M}_{m,n})$ that consists of bounded analytic matrix functions in the unit disk $\mathbb{D}$.

However, it is well known and it is easy to see that in the matrix case a continuous (and even infinitely smooth) function Φ generically has infinitely many best approximations by bounded analytic functions. It seems natural to impose additional constraints on a best approximation and choose among best approximations the "very best".

To introduce the notion of very best approximation, we recall that for a matrix A (or a Hilbert space operator A) the singular value $s_j(A)$, $j\ge 0$, is by definition the distance from A to the set of matrices (operators) of rank at most j. Clearly, $s_0(A)=\|A\|$.

Given a matrix function $\Phi\in L^\infty(\mathbb{M}_{m,n})$ we define inductively the sets $\boldsymbol{\Omega}_j$, $0\le j\le \min\{m,n\}-1$, by

$$\boldsymbol{\Omega}_0 = \{F\in H^\infty(\mathbb{M}_{m,n}):\ F \text{ minimizes } \ t_0 \stackrel{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} \|\Phi(\zeta)-F(\zeta)\|\};$$

$$\boldsymbol{\Omega}_j = \{F\in \Omega_{j-1}:\ F \text{ minimizes } \ t_j \stackrel{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} s_j(\Phi(\zeta)-F(\zeta))\}.$$

Functions in $\boldsymbol{\Omega}_{\min\{m,n\}-1}$ are called *superoptimal approximations* of Φ by bounded analytic matrix functions. The numbers $t_j=t_j(\Phi)$ are called the *superoptimal singular values* of Φ. Note that the functions in $\boldsymbol{\Omega}_0$ are just the best approximations by analytic matrix functions. The notion of superoptimal approximation was introduced in [Y].

It was proved in [PY1] that for a continuous $m\times n$ matrix function Φ there exists a unique superoptimal approximation F by bounded analytic matrix functions and the error function $\Phi-F$ satisfies

$$s_j(\Phi(\zeta)-F(\zeta)) = t_j(\Phi) \quad \text{almost everywhere on} \quad \mathbb{T}. \tag{1.1}$$

Later this uniqueness result was obtained in a different way by Treil [T].

In [PT] the uniqueness result was improved. It was shown in [PT] that if $\Phi\in L^\infty(\mathbb{M}_{m,n})$ and the essential norm $\|H_\Phi\|_{\mathrm{e}}$ of the Hankel operator H_Φ is less than

the smallest nonzero superoptimal singular value of Φ, then there exists a unique superoptimal approximation F by analytic matrix functions and (1.1) holds.

For $\Phi \in L^\infty(\mathbb{M}_{m,n})$ the Hankel operator $H_\Phi : H^2(\mathbb{C}^n) \to H^2_-(\mathbb{C}^n)$ is defined in the same way as in the scalar case: $H_\Phi f = \mathbb{P}_- \Phi f$, its norm is given by

$$\|H_\Phi\| = \operatorname{dist}_{L^\infty}(\Phi, H^\infty(\mathbb{M}_{m,n})),$$

([Pa]) and its essential norm is equal to

$$\|H_\Phi\|_{\mathrm{e}} = \operatorname{dist}_{L^\infty}(\Phi, (H^\infty + C)(\mathbb{M}_{m,n}))$$

(see e.g., [Sa] where the proof of this formula in the scalar case is given, in the matrix case the proof is the same). Clearly, $\|H_\Phi\|$ is equal to the largest superoptimal singular value $t_0(\Phi)$ of Φ.

In [PY1] and [PT] it was shown that if Φ satisfies the above conditions and F is the unique superoptimal approximation of Φ, then the error function $\Phi - F$ admits a so-called thematic factorization (see §2 for precise definitions). The technique of thematic factorizations turned out to be very fruitful (see [PY1], [PY2], [PT]). However, in the case of multiple superoptimal singular values the factors in a thematic factorization essentially depend on the choice of a factorization and by no means they can be determined by the matrix function Φ itself.

In this paper we consider a modification of the notion of a thematic factorization. Under the same assumptions on Φ we show that for the superoptimal approximation F the error function $\Phi - F$ admits a so-called canonical factorization. Unlike the case of thematic factorizations the factors in canonical factorizations are determined by the function Φ modulo constant unitary factors. Similarly, we consider so-called partial canonical factorizations for badly approximable matrix functions and prove the same invariance properties. This is done in §8.

Canonical factorizations are defined in terms of so called balanced unitary-valued matrix functions which are defined in §3. The notion of a balanced matrix function generalizes the notion of a thematic matrix function that was used to define a thematic factorization. We discuss some remarkable properties of balanced matrix functions in §3.

Recall that a matrix function Φ is called *badly approximable* if the zero function is a best approximation of Φ by bounded analytic matrix functions. A matrix function Φ is called *very badly approximable* if the zero function is a superoptimal approximation of Φ by bounded analytic matrix functions.

In §4 we prove that badly approximable matrix functions admit so-called partial canonical factorizations. In §5 and §6 we compare (partial) thematic factorizations with (partial) canonical factorizations and deduce a number of results on partial canonical factorizations from the corresponding results on partial thematic factorizations proved earlier in [PY1] and [PT]. Canonical factorizations of very badly approximable functions are discussed in §7.

§9 is devoted to hereditary properties of (partial) canonical factorizations. In other words for many important function spaces X we prove in §9 that if Φ is a badly approximable matrix function whose entries belong to X, then the entries of all factors in a partial canonical factorization of Φ also belong to X. In particular if $\Phi \in (H^\infty + C)(\mathbb{M}_{m,n})$, then the entries of all factors in a canonical factorization of Φ belong to the space QC of quasi-continuous functions.

In §10 we characterize the very badly approximable unitary-valued functions U under the assumption $\|H_U\|_{\mathrm{e}} < 1$. Such unitary-valued functions are involved in canonical factorizations.

Finally, in §2 we give definitions and state results that will be used in this paper.

2. **Preliminaries**

Toeplitz operators and Wiener–Hopf factorizations. For a matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$ the *Toeplitz operator* $T_\Phi : H^2(\mathbb{C}^n) \to H^2(\mathbb{C}^m)$ is defined by

$$T_\Phi f = \mathbb{P}_+ \Phi f, \quad f \in H^2(\mathbb{C}^n).$$

Suppose now that $m = n$. By Simonenko's theorem [Si] (see also [LS]), if T_Φ is Fredholm, then Φ admits a Wiener–Hopf factorization

$$\Phi = Q_2^* \begin{pmatrix} z^{d_1} & 0 & \cdots & 0 \\ 0 & z^{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z^{d_n} \end{pmatrix} Q_1^{-1},$$

where $d_1, \cdots, d_n \in \mathbb{Z}$, and Q_1 and Q_2 are matrix functions invertible in $H^2(\mathbb{M}_{n,n})$. It is always possible to arrange the *Wiener–Hopf indices* d_j is the nondecreasing order: $d_1 \le \cdots \le d_n$ in which case they are uniquely determined by Φ.

Maximizing vectors of vectorial Hankel operators. Let Φ be a matrix function in $L^\infty(\mathbb{M}_{m,n})$ such that the Hankel operator $H_\Phi : H^2(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$ has a maximizing vector $f \in H^2(\mathbb{C}^n)$. Let F be a best approximation of Φ by bounded analytic matrix functions. Put $g = H_\Phi f$. Then

$$H_\Phi f = (\Phi - F) f,$$

$$\|\Phi(\zeta) - F(\zeta)\|_{\mathbb{M}_{m,n}} = \|H_\Phi\| \quad \text{for almost all} \quad \zeta \in \mathbb{T},$$

$$\|g(\zeta)\|_{\mathbb{C}^m} = \|H_\Phi\| \cdot \|f(\zeta)\|_{\mathbb{C}^n} \quad \text{for almost all} \quad \zeta \in \mathbb{T},$$

and $f(\zeta)$ is a maximizing vector of $\Phi(\zeta) - F(\zeta)$ for almost all $\zeta \in \mathbb{T}$ ([AAK3], see also [PY1]).

Badly approximable scalar functions. If φ is a nonzero continuous scalar function on the unit circle, then φ is badly approximable if and only if φ has constant modulus and its winding number $\operatorname{wind}\varphi$ is negative (see [AAK1] and

[Po]). If φ is a nonzero scalar function in $H^\infty + C$, then φ is badly approximable if and only if $|\varphi|$ is constant almost everywhere on $\mathbb{T}$, $\varphi \in QC$ and $\operatorname{ind} T_\varphi > 0$ (see [PK]). Here

$$H^\infty + C \stackrel{\text{def}}{=} \{f + g : \ f \in C(\mathbb{T}), \ g \in H^\infty\}$$

is a closed subalgebra of L^∞ and

$$QC \stackrel{\text{def}}{=} \{f \in H^\infty + C : \ \bar{f} \in H^\infty + C\}.$$

Note that if $\varphi \in QC$ and $|\varphi| = \text{const} > 0$, then the Toeplitz operator T_φ is Fredholm.

These results can easily be generalized to the set of scalar functions $\varphi \in L^\infty$ satisfying $\|H_\varphi\|_{\mathrm{e}} < \|H_\varphi\|$. Under this condition φ is badly approximable if and only if $|\varphi| = \text{const}$ and $\operatorname{ind} T_\varphi > 0$. Again it is easy to see that if $|\varphi| = \text{const}$ and $\|H_\varphi\|_{\mathrm{e}} < \|H_\varphi\|$, then T_φ is Fredholm.

Inner and outer matrix functions. A matrix function $\Phi \in H^\infty(\mathbb{M}_{m,n})$ is called *inner* if $\Phi^*(\zeta)\Phi(\zeta) = I_n$ for almost all $\zeta \in \mathbb{T}$, where I_n stands for the identity $n \times n$ matrix (or the matrix function identically equal to I_n).

Consider the operator of multiplication by z on $H^2(\mathbb{C}^n)$. If $\mathcal{L}$ is a nonzero invariant subspace of this operator, then by the Beurling–Lax theorem, there exists an inner matrix function $\Theta \in H^\infty(\mathbb{M}_{n,r})$ such that

$$\mathcal{L} = \Theta H^2(\mathbb{C}^r).$$

In this case

$$\dim\{f(\zeta) : \ f \in \mathcal{L}\} = r \quad \text{for almost all} \quad \zeta \in \mathbb{D}.$$

If Θ° is an inner matrix function in $H^\infty(\mathbb{M}_{n,r^\circ})$ and $\Theta^\circ H^2(\mathbb{C}^{r^\circ}) = \Theta H^2(r)$, then $r = r^\circ$ and there exists a unitary matrix $\mathfrak{U} \in \mathbb{M}_{\mathfrak{r},\mathfrak{r}}$ such that $\Theta^\circ = \Theta\mathfrak{U}$.

A matrix function $F \in H^2(\mathbb{M}_{m,n})$ is called *outer* if the linear span of the set $\{Fz^j x : \ j \geq 0, \ x \in \mathbb{C}^n\}$ is dense in $H^2(\mathbb{C}^m)$. F is called *co-outer* if the transposed function F^{t} is outer.

If Ψ is a matrix function in $H^2(\mathbb{M}_{m,n})$, then there exist an inner matrix function Θ and an outer matrix function F such that $\Psi = \Theta F$. Moreover, if $\Psi = \Theta^\circ F^\circ$, where Θ° is inner and F° is outer, then there exists a unitary matrix $\mathfrak{U}$ such that $\Theta^\circ = \Theta\mathfrak{U}$ and $F^\circ = \mathfrak{U}^*\mathfrak{F}$.

The above results can be found in the books [SNF1] and [Ni].

Badly approximable matrix functions and thematic factorizations. Let $n \geq 2$. An $n \times n$ matrix function V is called *thematic* if it is unitary-valued and has the form

$$\begin{pmatrix} \boldsymbol{v} & \overline{\Theta} \end{pmatrix},$$

where $\boldsymbol{v}$ is an $n \times 1$ inner and co-outer column function and Θ is an $n \times (n-1)$ inner and co-outer matrix function. It is natural to say that a scalar function is thematic if it is constant and has modulus 1.

Let Φ be a matrix function in $L^\infty(\mathbb{M}_{m,n})$ such that $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. If F a best approximation of Φ, then $\Phi - F$ admits a factorization

$$\Phi - F = W^* \begin{pmatrix} t_0 u & 0 \\ 0 & \Psi \end{pmatrix} V^*, \tag{2.1}$$

where V and W^{t} are thematic matrix functions, u is a unimodular scalar function (i.e., $|u(\zeta)| = 1$ a.e. on $\mathbb{T}$) such that T_u is Fredholm and $\operatorname{ind} T_u > 0$, and Ψ is a matrix function in $L^\infty(\mathbb{M}_{m-1,n-1})$ such that $\|\Psi\|_{L^\infty} \le t_0$. This result was obtained in [PT]. Earlier the same fact was proved in [PY1] in the case $\Phi \in (H^\infty + C)(\mathbb{M}_{m,n})$. Moreover, it was shown in [PT] that $\|H_\Psi\|_{\mathrm{e}} \le \|H_\Phi\|_{\mathrm{e}}$ and it was proved in [PY1] that the problem of finding a superoptimal approximation of Φ reduces to the problem of finding a superoptimal approximation of Ψ.

Clearly, the left-hand side of (2.1) is a badly approximable function. Conversely, it follows from the results of [PY1] that if $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and Φ admits a factorization in the form

$$\Phi = W^* \begin{pmatrix} su & 0 \\ 0 & \Psi \end{pmatrix} V^*,$$

where V and W^{t} are thematic matrix functions, $s > 0$, u is a unimodular function such that T_u is Fredholm and $\operatorname{ind} T_u > 0$, and $\|\Psi\|_{L^\infty} \le s$, then Φ is badly approximable and $s = t_0(\Phi) = \|H_\Phi\|$.

Suppose now that $\|H_\Phi\|_{\mathrm{e}} < t_1$. The inequality $\|H_\Psi\|_{\mathrm{e}} \le \|H_\Phi\|_{\mathrm{e}}$ proved in [PT] allows one to continue the diagonalization process and prove that if $l \le \min\{m, n\}$, $\|H_\Phi\|_{\mathrm{e}} < t_{l-1}$ and $t_{l-1} > t_l$, then for any matrix function $F \in \mathbf{\Omega}_{l-1}$ (the sets $\mathbf{\Omega}_j$ are defined in §1) the matrix function $\Phi - F$ admits a factorization

$$\Phi - F = W_0^* \cdots W_{l-1}^* \begin{pmatrix} t_0 u_0 & 0 & \cdots & 0 & 0 \\ 0 & t_1 u_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & t_{l-1} u_{l-1} & 0 \\ 0 & 0 & \cdots & 0 & \Psi \end{pmatrix} V_{l-1}^* \cdots V_0^*, \tag{2.2}$$

where

$$W_j = \begin{pmatrix} I_j & 0 \\ 0 & \breve{W}_j \end{pmatrix}, \quad V_j = \begin{pmatrix} I_j & 0 \\ 0 & \breve{V}_j \end{pmatrix}, \quad 1 \le j \le l-1,$$

the $W_0^{\mathrm{t}}, \breve{W}_j^{\mathrm{t}}, V_0, \breve{V}_j$ are thematic matrix functions, the u_j are unimodular functions such that T_{u_j} is Fredholm and $\operatorname{ind} T_{u_j} > 0$, $\|\Psi\|_{L^\infty} \le t_{l-1}$, and $\|H_\Psi\| < t_{l-1}$. Factorizations of the form (2.2) are called *partial thematic factorizations*.

Finally, if $\|H_\Phi\|_{\mathrm{e}}$ is less than the smallest nonzero superoptimal singular value of Φ and F is the unique superoptimal approximation of Φ, then $\Phi - F$ admits a

factorization

$$\Phi - F = W_0^* \cdots W_{\iota-1}^* \begin{pmatrix} t_0 u_0 & 0 & \cdots & 0 & 0 \\ 0 & t_1 u_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & t_{\iota-1} u_{\iota-1} & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} V_{\iota-1}^* \cdots V_0^*, \tag{2.3}$$

where the u_j, V_j, W_j are as above and $t_{\iota-1}$ is the smallest nonzero superoptimal singular value of Φ. Factorizations of the form (2.3) are called *thematic factorizations*. This result was obtained in [PT]. Earlier the same fact was proved in [PY1] in the case $\Phi \in (H^\infty + C)(\mathbb{M}_{m,n})$. Note that the lower right entry of the diagonal matrix function on the right-hand side of (2.3) has size $(m-\iota) \times (n-\iota)$ and the numbers $m-\iota$ or $n-\iota$ may be zero.

Clearly, the left-hand side of (2.3) is a very badly approximable function. Conversely, it follows from the results of [PY1] that if a matrix function admits a thematic factorization, it is very badly approximable.

A disadvantage of thematic factorizations is that a thematic factorization may essentially depend on the choice of matrix functions V_j and W_j and they are not uniquely determined by the given matrix function. Moreover, with any nonzero superoptimal singular value $t_j(\Phi)$ we can associate the *factorization index* $k_j \stackrel{\text{def}}{=} \operatorname{ind} T_{u_j}$ of a thematic (or partial thematic) factorization. It was shown in [PY1] that in the case of multiple superoptimal singular values even the indices k_j are not uniquely determined by Φ. We refer the reader to [PY2], [PT], and [AP1] for further results on the thematic factorization indices k_j. In particular, in [AP1] it was proved that it is always possible to choose a so-called monotone (partial) thematic factorization and the indices of a monotone (partial) thematic factorization are uniquely determined by Φ. However, the matrix function V_j and W_j are still essentially dependent of our choice.

That is why we introduce in this paper (partial) canonical factorizations and we prove that they are "essentially" unique.

3. **Balanced unitary-valued matrix functions**

In [PY1] and [PT] thematic matrix functions and thematic factorizations played a crucial role to study superoptimal approximation. We consider here a more general class of balanced unitary-valued matrix functions.

Definition. Let n be a positive integer and let r be an integer such that $0 < r < n$. Suppose that Υ is an $n \times r$ inner and co-outer matrix function and Θ

is an $n \times (n-r)$ inner and co-outer matrix function. If the matrix function

$$\mathcal{V} = \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}$$

is unitary-valued, it is called an *r-balanced matrix function.* If $r = 0$ or $r = n$, it is natural to say that an r-balanced matrix is a constant unitary matrix. An $n \times n$ matrix function $\mathcal{V}$ is called *balanced* if it is r-balanced for some r, $0 \le r \le n$.

Recall that *thematic* matrix functions are just 1-balanced according to this definition.

It is known (see [V]) that if $0 < r < n$ and Υ is an $n \times r$ inner and co-outer matrix function, then it has a balanced unitary completion, i.e., there exists an $n \times (n-r)$ matrix function Θ such that $\begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}$ is balanced. Moreover, such a completion is unique modulo a right constant unitary factor. We are going to study some interesting properties of balanced matrix functions and we need a construction of the complementary matrix function Θ. That is why we give the construction here.

We will see that balanced matrix functions have many nice properties which can justify the term "balanced".

Note that the case $r = 1$ was studied in [PY1]. However, it turns out that studying the more general case of an arbitrary r simplifies the approach given in [PY1].

Theorem 3.1. *Let n be a positive integer, $0 < r < n$, and let Υ be and $n \times r$ inner and co-outer matrix function. Then the subspace $\mathcal{L} \stackrel{\text{def}}{=} \operatorname{Ker} T_{\Upsilon^{\rm t}}$ has the form*

$$\mathcal{L} = \Theta H^2(\mathbb{C}^{n-r}),$$

where Θ is an inner and co-outer $n \times (n-r)$ matrix function such that $\begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}$ is balanced.

Proof. Clearly, the subspace $\mathcal{L}$ of $H^2(\mathbb{C}^n)$ is invariant under multiplication by z. By the Beurling–Lax theorem (see §2), $\mathcal{L} = \Theta H^2(\mathbb{C}^l)$ for some $l \le n$ and an $n \times l$ inner matrix function Θ.

Let us first prove that Θ is co-outer. Suppose that $\Theta^{\rm t} = \mathcal{O}\mathcal{F}$ where $\mathcal{O}$ is an inner matrix function and F is an outer matrix function. Since Θ is inner, it is easy to see that $\mathcal{O}$ has size $l \times l$ while F has size $l \times n$. It follows that $\mathcal{O}^*\Theta^{\rm t} = F$, and so $F^{\rm t} = \Theta\overline{\mathcal{O}}$. Since both Θ and $\overline{\mathcal{O}}$ take isometric values almost everywhere on $\mathbb{T}$, the matrix function $F^{\rm t}$ is inner.

Let us show that

$$F^{\rm t} H^2(\mathbb{C}^l) \subset \mathcal{L}. \tag{3.1}$$

First of all, it is easy to see that $\Upsilon^{\rm t}\Theta$ is the zero matrix function. Let now $f \in H^2(\mathbb{C}^l)$. We have

$$\Upsilon^{\rm t} F^{\rm t} f = \Upsilon^{\rm t}\Theta\overline{\mathcal{O}} f = 0,$$

which proves (3.1).

It follows from (3.1) that

$$F^{\mathrm{t}} H^2(\mathbb{C}^l) = \Theta \overline{\mathcal{O}} H^2(\mathbb{C}^l) \subset \Theta H^2(\mathbb{C}^l).$$

Multiplying the last inclusion by Θ^*, we have

$$\overline{\mathcal{O}} H^2(\mathbb{C}^l) \subset H^2(\mathbb{C}^l),$$

which implies that $\mathcal{O}$ is a constant unitary matrix function.

Let us now prove that $l = n - r$. First of all it is evident that the columns of $\Upsilon(\zeta)$ are orthogonal to the columns of $\overline{\Theta}(\zeta)$ almost everywhere on $\mathbb{T}$. Hence, the matrix function $\begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}$ takes isometric values almost everywhere, and so $l \le n - r$.

To show that $l \ge n - r$, consider the functions $P_{\mathcal{L}} C$, where $P_{\mathcal{L}}$ is the orthogonal projection onto $\mathcal{L}$ and C is a constant function which we identify with a vector in $\mathbb{C}^n$. Note that $C \perp \mathcal{L}$ if and only if the vectors $f(0)$ and C are orthogonal in $\mathbb{C}^n$ for any $f \in \mathcal{L}$. Let us prove that

$$\dim\{f(0) :\ f \in \mathcal{L}\} \ge n - r. \tag{3.2}$$

Since Υ is co-outer, it is easy to see that $\operatorname{rank} \Upsilon(0) = r$. Without loss of generality we may assume that $\Upsilon = \begin{pmatrix} \Upsilon_1 \\ \Upsilon_2 \end{pmatrix}$, where Υ_1 has size $(n-r) \times r$, Υ_2 has size $r \times r$, and the matrix $\Upsilon_2(0)$ is invertible. Let now K be an arbitrary vector in $\mathbb{C}^{n-r}$. Put

$$f = (\det \Upsilon_2(0))^{-1} \det \Upsilon_2 \begin{pmatrix} K \\ -(\Upsilon_2^{\mathrm{t}})^{-1} \Upsilon_1^{\mathrm{t}} K \end{pmatrix}.$$

Clearly, $f \in H^2(\mathbb{C}^n)$. We have

$$\Upsilon^{\mathrm{t}} f = \begin{pmatrix} \Upsilon_1^{\mathrm{t}} & \Upsilon_2^{\mathrm{t}} \end{pmatrix} f = (\det \Upsilon_2(0))^{-1} \det \Upsilon_2 (\Upsilon_1^{\mathrm{t}} K - \Upsilon_2^{\mathrm{t}} (\Upsilon_2^{\mathrm{t}})^{-1} \Upsilon_1^{\mathrm{t}} K) = 0,$$

and so $f \in \mathcal{L}$. On the other hand, it is easy to see that

$$f(0) = \begin{pmatrix} K \\ -(\Upsilon_2^{\mathrm{t}}(0))^{-1} \Upsilon_1^{\mathrm{t}}(0) K \end{pmatrix}$$

and since K is an arbitrary vector in $\mathbb{C}^{n-r}$, this proves (3.2).

We have already observed that

$$\{f(0) :\ f \in \mathcal{L}\} = \mathbb{C}^n \ominus \{C \in \mathbb{C}^n :\ P_{\mathcal{L}} C = 0\},$$

and so it follows from (3.2) that

$$\dim\{P_{\mathcal{L}} C :\ C \in \mathbb{C}^n\} \ge n - r.$$

It is easy to see that for $C \in \mathbb{C}^n$ we have

$$P_{\mathcal{L}} C = \Theta \mathbb{P}_+ \Theta^* C = \Theta(\Theta^*(0)) C.$$

Clearly, $\Theta(\Theta^*(0))C$ belongs to the linear span of the columns of Θ. This completes the proof of the fact that $l = n - r$ and proves that $\begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}$ is a balanced matrix function. ■

It is well known (see [V]) that given an inner and co-outer matrix function Υ, the balanced completion Θ is unique modulo a right constant unitary factor. Indeed, if $\begin{pmatrix} \Upsilon & \overline{\Theta}_1 \end{pmatrix}$ is another balanced matrix function, then clearly

$$\Theta_1 H^2(\mathbb{C}^{n-r}) \subset \mathcal{L} = \Theta H^2(\mathbb{C}^{n-r}).$$

It follows (see [Ni]) that $\Theta_1 = \Theta\mathcal{O}$ for an inner matrix function $\mathcal{O}$. Clearly, $\mathcal{O}$ has size $(n-r) \times (n-r)$. Hence, $\Theta_1^{\mathrm{t}} = \mathcal{O}^{\mathrm{t}}\Theta^{\mathrm{t}}$, $\mathcal{O}^{\mathrm{t}}$ is inner, and since Θ_1 is co-outer, it follows that $\mathcal{O}$ is a unitary constant.

Next, we are going to study the property of analyticity of minors of balanced matrix functions. Let $\mathcal{V} = \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}$ be an r-balanced $n \times n$ matrix function. We are going to study its minors $\mathcal{V}_{\imath_1 \cdots \imath_k, \jmath_1 \cdots \jmath_k}$ of order k, i.e., the determinants of the submatrix of $\mathcal{V}$ with rows $\imath_1, \cdots, \imath_k$ and columns $\jmath_1, \cdots, \jmath_k$. Here $1 \le \imath_1 < \cdots < \imath_k \le n$ and $1 \le \jmath_1 < \cdots < \jmath_k \le n$. By a *minor of $\mathcal{V}$ on the first r columns* we mean a minor $\mathcal{V}_{\imath_1 \cdots \imath_k, \jmath_1 \cdots \jmath_k}$ with $k \ge r$ and $\jmath_1 = 1, \cdots, \jmath_r = r$. Similarly, by a *minor of $\mathcal{V}$ on the last $n - r$ columns* we mean a minor $\mathcal{V}_{\imath_1 \cdots \imath_k, \jmath_1 \cdots \jmath_k}$ with $k \ge n - r$ and $\jmath_{k-n+r+1} = r + 1, \cdots, \jmath_k = n$.

It was proved in [PY1] (Theorem 1.1) that for a thematic (i.e., 1-balanced) matrix function all minors on the first column are in H^∞. The following result generalizes that theorem and makes it more symmetric.

Theorem 3.2. *Let $\mathcal{V}$ be an r-balanced matrix function of size $n \times n$. Then all minors of $\mathcal{V}$ on the first r columns are in H^∞ while all minors of $\mathcal{V}$ on the last $n - r$ columns are in $\overline{H^\infty}$.*

Proof. It is easy to see that it is sufficient to prove the theorem for the minors on the first r columns of $\mathcal{V}$. To deduce the second assertion of the theorem, we can consider the matrix function $\overline{\mathcal{V}}$ and rearrange its columns to make it $(n-r)$-balanced.

Denote by $\Upsilon_1, \cdots, \Upsilon_r$ and $\Theta_1, \cdots, \Theta_{n-r}$ the columns of Υ and Θ. In the proof of Theorem 3.1 we have observed that for any constant $C \in \mathbb{C}^n$ we have

$$P_{\mathcal{L}}C = \Theta\Theta^*(0)C$$

and

$$\dim\{P_{\mathcal{L}}C : \ C \in \mathbb{C}^n\} = \dim\{\Theta^*(0)C : \ C \in \mathbb{C}^n\} = n - r.$$

It follows that there exist $C_1, \cdots, C_{n-r} \in \mathbb{C}^n$ such that

$$\Theta_j = P_{\mathcal{L}}C_j, \quad 1 \le j \le n - r. \tag{3.3}$$

If $1 \le d \le n - r$ and $1 \le j_1 < j_2 < \cdots < j_d \le n - r$, we consider the vector function

$$\Upsilon_1 \wedge \cdots \wedge \Upsilon_r \wedge \overline{\Theta}_{j_1} \wedge \cdots \wedge \overline{\Theta}_{j_d}$$

whose $\begin{pmatrix} n \\ r+d \end{pmatrix}$ components are the minors of order $r+d$ of the matrix function $\begin{pmatrix} \Upsilon_1 & \cdots & \Upsilon_r & \overline{\Theta}_{j_1} & \cdots & \overline{\Theta}_{j_d} \end{pmatrix}$.

It follows from (3.3) that

$$\Theta_j = C_j - P_{\mathcal{L}^\perp} C_j,$$

where $P_{\mathcal{L}^\perp}$ is the orthogonal projection onto $\mathcal{L}^\perp = \operatorname{clos}\operatorname{Range} T_{\overline{\Upsilon}}$. We have

$$\Upsilon_1 \wedge \cdots \wedge \Upsilon_r \wedge \overline{\Theta}_{j_1} \wedge \cdots \wedge \overline{\Theta}_{j_d} = \Upsilon_1 \wedge \cdots \wedge \Upsilon_r \wedge (\overline{C}_{j_1} - \overline{P_{\mathcal{L}^\perp} C_{j_1}}) \wedge \cdots \wedge (\overline{C}_{j_d} - \overline{P_{\mathcal{L}^\perp} C_{j_d}}).$$

The components of this vector belong to L^∞ and can be approximated in $L^{2/d}$ by vector functions of the form

$$\Upsilon_1 \wedge \cdots \wedge \Upsilon_r \wedge (\overline{C}_{j_1} - \overline{g}_{j_1}) \wedge \cdots \wedge (\overline{C}_{j_d} - \overline{g}_{j_d}), \tag{3.4}$$

where $g_{j_1}, \cdots, g_{j_d} \in \operatorname{Range} T_{\overline{\Upsilon}}$. Hence, it is sufficient to prove that the components of (3.4) belong to $H^{2/d}$. Let $g_{j_l} = \mathbb{P}_+ \overline{\Upsilon} f_l$ for $f_l \in H^2(\mathbb{C}^r)$, $1 \le l \le d$. We have

$$\begin{aligned} & \Upsilon_1 \wedge \cdots \wedge \Upsilon_r \wedge (\overline{C}_{j_1} - \overline{g}_{j_1}) \wedge \cdots \wedge (\overline{C}_{j_d} - \overline{g}_{j_d}) \\ = \; & \Upsilon_1 \wedge \cdots \wedge \Upsilon_r \wedge (\overline{C}_{j_1} - \overline{\mathbb{P}_+ \overline{\Upsilon} f_1}) \wedge \cdots \wedge (\overline{C}_{j_d} - \overline{\mathbb{P}_+ \overline{\Upsilon} f_d}) \\ = \; & \Upsilon_1 \wedge \cdots \wedge \Upsilon_r \wedge (\overline{C}_{j_1} - \Upsilon \bar{f}_1 + \overline{\mathbb{P}_- \overline{\Upsilon} f_1}) \wedge \cdots \wedge (\overline{C}_{j_d} - \Upsilon \bar{f}_d + \overline{\mathbb{P}_- \overline{\Upsilon} f_d}). \end{aligned}$$

Clearly, almost everywhere on $\mathbb{T}$ the vectors $\Upsilon(\zeta)\overline{f_l(\zeta)}$ are linear combinations of $\Upsilon_1(\zeta), \cdots, \Upsilon_r(\zeta)$. Therefore if we expand the above wedge product using the multilinearity of $\wedge$, all terms containing $\Upsilon \bar{f}_l$, give zero contribution. Thus we have

$$\begin{aligned} & \Upsilon_1 \wedge \cdots \wedge \Upsilon_r \wedge (\overline{C}_{j_1} - \Upsilon \bar{f}_1 + \overline{\mathbb{P}_- \overline{\Upsilon} f_1}) \wedge \cdots \wedge (\overline{C}_{j_d} - \Upsilon \bar{f}_d + \overline{\mathbb{P}_- \overline{\Upsilon} f_d}) \\ = \; & \Upsilon_1 \wedge \cdots \wedge \Upsilon_r \wedge (\overline{C}_{j_1} + \overline{\mathbb{P}_- \overline{\Upsilon} f_1}) \wedge \cdots \wedge (\overline{C}_{j_d} + \overline{\mathbb{P}_- \overline{\Upsilon} f_d}) \in H^{2/d}. \quad \blacksquare \end{aligned}$$

The following immediate consequence of Theorem 3.2 was obtained in [PY1] (Theorem 1.2) for $r = 1$ by another method.

Corollary 3.3. *Let $\mathcal{V}$ be a balanced matrix function. Then* $\det \mathcal{V}$ *is a constant function of modulus 1.*

We shall need another nice property of balanced matrix functions that was obtained in [Pe4]. Namely, it was proved in Lemma 6.2 of [Pe4] that for a balanced matrix function $\mathcal{V}$ of size $n \times n$ the Toeplitz operator $T_{\mathcal{V}} : H^2(\mathbb{C}^n) \to H^2(\mathbb{C}^n)$ has trivial kernel and dense range. If we apply that result to $\overline{\mathcal{V}}$ with rearranged columns, we see that the Toeplitz operator $T_{\overline{\mathcal{V}}}$ also has trivial kernel and dense range.

Finally, we obtain in this section an analog of Lemma 1.5 of [PY1] where the case of thematic matrix functions was considered.

Theorem 3.4. *Let $0 < r \le \min\{m,n\}$ and let $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$ be r-balanced matrix functions of sizes $n \times n$ and $m \times m$ respectively. Then*

$$\mathcal{W}H^\infty(\mathbb{M}_{m,n})\mathcal{V} \bigcap \begin{pmatrix} 0 & 0 \\ 0 & L^\infty(\mathbb{M}_{m-r,n-r}) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & H^\infty(\mathbb{M}_{m-r,n-r}) \end{pmatrix}.$$

The proof of Theorem 3.4 is exactly the same as the proof of its special case Lemma 1.5 of [PY1].

4. **Best approximation and partial canonical factorizations**

For a matrix function Φ in $L^\infty(\mathbb{M}_{m,n})$ satisfying the condition $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$ and a best approximation F of Φ by bounded analytic matrix functions we obtain a so-called partial canonical factorization of $\Phi - F$. We characterize badly approximable functions satisfying this condition in terms of partial canonical factorizations. To this end we begin this section with the study of the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^n)$ that contains all maximizing vectors of H_Φ.

Theorem 4.1. *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. Let $\mathcal{M}$ be the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^n)$ that contains all maximizing vectors of H_Φ. Then*

$$\mathcal{M} = \Upsilon H^2(\mathbb{C}^r), \tag{4.1}$$

where r is the number of superoptimal singular values of Φ equal to $\|H_\Phi\|$, Υ is an inner and co-outer $n \times r$ matrix function.

Proof. Consider first the case $m = n$. Without loss of generality we may assume that $\|H_\Phi\| = 1$. It follows from the results of §3 of [AP2] that there exists a unitary interpolant $\mathcal{U}$ of Φ (i.e., a unitary-valued matrix function $\mathcal{U}$ satisfying $\hat{\mathcal{U}}(j) = \hat{\Phi}(j)$, $j < 0$) such that the Toeplitz operator $T_{\mathcal{U}}$ is Fredholm and each such unitary interpolant has precisely r negative Wiener–Hopf indices. Consider a Wiener–Hopf factorization of $\mathcal{U}$

$$\mathcal{U} = Q_2^* \begin{pmatrix} z^{d_1} & 0 & \cdots & 0 \\ 0 & z^{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z^{d_n} \end{pmatrix} Q_1^{-1}, \tag{4.2}$$

where Q_1 and Q_2 are matrix functions invertible in $H^2(\mathbb{M}_{n,n})$, and $d_1 \le d_2 \cdots \le d_n$. Since $\mathcal{U}$ has r negative Wiener–Hopf indices, we have

$$d_1 \le \cdots \le d_r < 0 \le d_{r+1} \le \cdots \le d_n.$$

Clearly, $H_\Phi = H_{\mathcal{U}}$. It is also easy to see that a nonzero function $f \in H^2(\mathbb{C}^n)$ is a maximizing vector of H_Φ if an only if $f \in \operatorname{Ker} T_{\mathcal{U}}$. It is well known and it is easy to see from (4.2) that

$$\operatorname{Ker} T_{\mathcal{U}} = \left\{ Q_1 \begin{pmatrix} q_1 \\ \vdots \\ q_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} : \ q_j \in \boldsymbol{\mathcal{P}}_+, \ \deg q_j < -d_j, \ 1 \le j \le r \right\}. \tag{4.3}$$

Here we denote by $\boldsymbol{\mathcal{P}}_+$ the set of analytic polynomials. Since $\mathcal{M}$ is the minimal invariant subspace of multiplication by z that contains $\operatorname{Ker} T_{\mathcal{U}}$, it follows from (4.3) that

$$\mathcal{M} = \operatorname{clos}_{H^2(\mathbb{C}^n)} \left\{ Q_1 \begin{pmatrix} q_1 \\ \vdots \\ q_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} : \ q_j \in \boldsymbol{\mathcal{P}}_+, \ 1 \le j \le r \right\}. \tag{4.4}$$

Since $Q_1(\zeta)$ is an invertible matrix for all $\zeta \in \mathbb{D}$, it follows easily from (4.4) that $\dim\{f(\zeta) : \ f \in \mathcal{M}\} = r$ for all $\zeta \in \mathbb{D}$. Therefore the z-invariant subspace $\mathcal{M}$ has the form $\mathcal{M} = \Upsilon H^2(\mathbb{C}^r)$, where Υ is an $n \times r$ inner matrix function (see [Ni]). It remains to prove that Υ is co-outer.

Denote by $Q_\heartsuit$ the matrix function obtained from Q_1 by deleting the last $n - r$ columns. It is easy to see that Υ is an inner part of $Q_\heartsuit$. Let $Q_\heartsuit = \Upsilon F$, where F is an $r \times r$ outer matrix function.

Denote by $Q_\spadesuit$ the matrix function obtained from Q_1^{-1} by deleting the last $n-r$ rows. Clearly, $Q_\spadesuit(\zeta)Q_\heartsuit(\zeta) = I_r$ for almost all $\zeta \in \mathbb{T}$.

We have

$$I_r = Q_\spadesuit Q_\heartsuit = Q_\spadesuit \Upsilon F,$$

and so

$$I_r = F^{\rm t} \Upsilon^{\rm t} Q_\spadesuit^{\rm t}.$$

Both $F^{\rm t}$ and $\Upsilon^{\rm t} Q_\spadesuit^{\rm t}$ are $r \times r$ matrix functions. Hence,

$$I_r = \Upsilon^{\rm t} Q_\spadesuit^{\rm t} F^{\rm t}.$$

It follows that $\Upsilon^{\rm t}$ is outer, and so Υ is co-outer.

Consider now the case $m < n$. Let $\Phi_\#$ be the matrix function obtained from Φ by adding $n-m$ zero rows. It is easy to see that the Hankel operators H_Φ and $H_{\Phi_\#}$ have the same maximizing vectors. This reduces the problem to the case $m = n$.

Finally, assume that $m > n$. Let $\Phi_\flat$ be the matrix function obtained from Φ by adding $m-n$ zero columns. It is easy to see that f is a maximizing vector of $H_{\Phi_\flat}$ if and only if it can be obtained from a maximizing vector of H_Φ by adding $m-n$ zero coordinates. Let $\mathcal{M}_\flat$ be the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^m)$ that contains all maximizing vectors of $H_{\Phi_\flat}$. Clearly, the number of superoptimal singular values of $\Phi_\flat$ equal to 1 is still r. Therefore there exists an $m \times r$ inner and co-outer matrix function $\Upsilon_\flat$ such that

$$\mathcal{M}_\flat = \Upsilon_\flat H^2(\mathbb{C}^r).$$

It is easy to see that the last $m-n$ rows of $\Upsilon_\flat$ are zero. Denote by Υ the matrix function obtained from $\Upsilon_\flat$ by deleting the last $m-n$ zero rows. Obviously, Υ is an inner and co-outer $n \times r$ matrix function and $\mathcal{M} = \Upsilon H^2(\mathbb{C}^r)$. ■

We need the following result.

Lemma 4.2. *Suppose that Φ satisfies the hypotheses of Theorem 4.1 and $\mathcal{M}$ is given by* (4.1)*. If $\|\Phi\|_{L^\infty} = \|H_\Phi\|$ and f is a nonzero vector function in $\mathcal{M}$, then $f(\zeta)$ is a maximizing vector of $\Phi(\zeta)$ for almost all $\zeta \in \mathbb{T}$.*

Proof. As we have noted in §2, if f is a maximizing vector of H_Φ, then $f(\zeta)$ is a maximizing vector of Φ for almost all $\zeta \in \mathbb{T}$ and $\|\Phi(\zeta)\|_{\mathbb{M}_{m,n}} = \|H_\Phi\|$ almost everywhere. Without loss of generality we may assume that $\|H_\Phi\| = 1$.

Let L be the set of vector functions of the form

$$q_1 g_1 + \cdots + q_M g_M,$$

where $q_j \in \boldsymbol{\mathcal{P}}_+$ and the g_j are maximizing vectors of H_Φ. By definition, $\mathcal{M}$ is the norm closure of L. Since the $g_j(\zeta)$ are maximizing vectors of $\Phi(\zeta)$ for almost all $\zeta \in \mathbb{T}$ (see §2), it follows that for $g \in L$, $g(\zeta)$ is a maximizing vector of $\Phi(\zeta)$ almost everywhere on $\mathbb{T}$. Let $\{f_j\}$ be a sequence of vector functions in L that converges to $f \in \mathcal{M}$ in $H^2(\mathbb{C}^n)$. Clearly,

$$\begin{aligned}\int_{\mathbb{T}} \|\Phi(\zeta)f(\zeta)\|^2_{\mathbb{C}^m} d\boldsymbol{m}(\zeta) &= \lim_{j\to\infty} \int_{\mathbb{T}} \|\Phi(\zeta)f_j(\zeta)\|^2_{\mathbb{C}^m} d\boldsymbol{m}(\zeta) \\ &= \lim_{j\to\infty} \int_{\mathbb{T}} \|f_j(\zeta)\|^2_{\mathbb{C}^m} d\boldsymbol{m}(\zeta) \\ &= \int_{\mathbb{T}} \|f(\zeta)\|^2_{\mathbb{C}^m} d\boldsymbol{m}(\zeta),\end{aligned}$$

and since obviously, $\|\Phi(\zeta)f(\zeta)\|_{\mathbb{C}^m} \le \|f(\zeta)\|_{\mathbb{C}^m}$ almost everywhere on $\mathbb{T}$, it follows that $f(\zeta)$ is a maximizing vector of $\Phi(\zeta)$ for almost all $\zeta \in \mathbb{T}$. ■

Again, suppose that Φ is as in Theorem 4.1. Obviously, a vector function f in $H^2(\mathbb{C}^n)$ is a maximizing vector of H_Φ if and only if $g \stackrel{\text{def}}{=} H_\Phi f \in H^2_-(\mathbb{C}^m)$ is a maximizing vector of H^*_Φ. Let us show that a vector function g in $H^2_-(\mathbb{C}^m)$ is a maximizing vector of H^*_Φ if and only if $\bar{z}\bar{g} \in H^2(\mathbb{C}^m)$ is a maximizing vector of

$H_{\Phi^{\mathrm{t}}}$. Indeed, assume that $\|H_\Phi\| = 1$. Then g is a maximizing vector of H_Φ^* if and only if

$$\|H_\Phi^* g\|_2 = \|\mathbb{P}_+ \Phi^* g\|_2 = \|g\|_2.$$

Clearly, this is equivalent to the equality

$$\|\mathbb{P}_- \Phi^{\mathrm{t}} \bar{z}\bar{g}\|_2 = \|\bar{z}\bar{g}\|_2,$$

which means that $\bar{z}\bar{g}$ is a maximizing vector of $H_{\Phi^{\mathrm{t}}}$.

It is easy to see that the matrix functions Φ and Φ^{t} have the same superoptimal singular values. Let $\mathcal{N}$ be the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^m)$ that contains all maximizing vectors of $H_{\Phi^{\mathrm{t}}}$. By Theorem 4.1, there exists an inner and co-outer matrix function $\Omega \in H^\infty(\mathbb{M}_{m,r})$ such that

$$\mathcal{N} = \Omega \mathcal{H}^{\in}(\mathbb{C}^{\nabla}).$$

By Theorem 3.1, Υ and Ω have balanced completions, i.e., there exist inner and co-outer matrix functions $\Theta \in H^\infty(\mathbb{M}_{n,n-r})$ and $\Xi \in H^\infty(\mathbb{M}_{m,m-r})$ such that

$$\mathcal{V} \stackrel{\text{def}}{=} \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix} \quad \text{and} \quad \mathcal{W}^{\mathrm{t}} \stackrel{\text{def}}{=} \begin{pmatrix} \Omega & \Xi \end{pmatrix} \tag{4.5}$$

are unitary-valued matrix functions.

Theorem 4.3. *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and $\|H_\Phi\|_{\mathrm{e}} < t_0 = \|H_\Phi\|$. Let r be the number of superoptimal singular values of Φ equal to t_0. Suppose that F is a best approximation of Φ by analytic matrix functions. Then $\Phi - F$ admits a factorization of the form*

$$\Phi - F = \mathcal{W}^* \begin{pmatrix} t_0 U & 0 \\ 0 & \Psi \end{pmatrix} \mathcal{V}^*, \tag{4.6}$$

where $\mathcal{V}$ and $\mathcal{W}$ are given by (4.5)*, U is an $r \times r$ unitary-valued very badly approximable matrix function such that $\|H_U\|_{\mathrm{e}} < 1$, and Ψ is a matrix-function in $L^\infty(\mathbb{M}_{m-r,n-r})$ such that $\|\Psi\|_{L^\infty} \le t_0$ and $\|H_\Psi\| = t_r(\Phi) < \|H_\Phi\|$. Moreover, U is uniquely determined by the choice of Υ and Ω, and does not depend on the choice of F.*

Proof. Without loss of generality we may assume that $\|H_\Phi\| = 1$. It follows from Lemma 4.2 that the columns of $\Upsilon(\zeta)$ are maximizing vectors of $\Phi(\zeta) - F(\zeta)$ for almost all $\zeta \in \mathbb{T}$. Similarly, the columns of $\Omega(\zeta)$ are maximizing vectors of $\Phi^{\mathrm{t}}(\zeta) - F^{\mathrm{t}}(\zeta)$ almost everywhere on $\mathbb{T}$.

We need two elementary lemmas.

Lemma 4.4. *Let $A \in \mathbb{M}_{m,n}$ and $\|A\| = 1$. Suppose that $v_1, \cdots, v_r$ is an orthonormal family of maximizing vectors of A and $w_1, \cdots, w_r$ is an orthonormal family of maximizing vectors of A^t. Then*

$$\begin{pmatrix} w_1 & \cdots & w_r \end{pmatrix}^{\mathrm{t}} A \begin{pmatrix} v_1 & \cdots & v_r \end{pmatrix}$$

is a unitary matrix.

Lemma 4.5. *Let A be a matrix in $\mathbb{M}_{m,n}$ such that $\|A\| = 1$ and A has the form*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where A_{11} is a unitary matrix. Then A_{12} and A_{21} are the zero matrices.

Both lemmas are obvious. Let us complete the proof of Theorem 4.3.

Consider the matrix function

$$\begin{pmatrix} U & X \\ Y & \Psi \end{pmatrix} \stackrel{\text{def}}{=} \mathcal{W}(\Phi - F)\mathcal{V}.$$

Here $U \in L^\infty(\mathbb{M}_{r,r})$, $X \in L^\infty(\mathbb{M}_{r,n-r})$, $Y \in L^\infty(\mathbb{M}_{m-r,r})$, and $\Psi \in L^\infty(\mathbb{M}_{m-r,n-r})$. If we apply Lemma 4.4 to the matrices $(\Phi - F)(\zeta)$, $\zeta \in \mathbb{T}$, and the columns of $\Upsilon(\zeta)$ and $\Omega(\zeta)$, we see $U = \Omega^{\rm t}(\Phi - F)\Upsilon$ is unitary-valued. By Lemma 4.5, X and Y are the zero matrix functions which proves (4.6).

Let us prove that $\|H_U\|_{\rm e} < 1$. Since

$$U = \Omega^{\rm t}(\Phi - F)\Upsilon, \tag{4.7}$$

we have

$$\begin{aligned} \|H_U\|_{\rm e} &= \operatorname{dist}_{L^\infty}\big(U, (H^\infty + C)(\mathbb{M}_{r,r})\big) \\ &= \operatorname{dist}_{L^\infty}\big(\Omega^{\rm t}\Phi\Upsilon, (H^\infty + C)(\mathbb{M}_{r,r})\big) \\ &\le \operatorname{dist}_{L^\infty}\big(\Phi, (H^\infty + C)(\mathbb{M}_{m,n})\big) = \|H_\Phi\|_{\rm e} < 1. \end{aligned}$$

Now it is turn to show that U is very badly approximable. Denote by $\mathcal{L}$ the minimal invariant subspace of multiplication by z that contains all maximizing vectors of H_U. Suppose that f is a maximizing vector of H_Φ. Then $f = \Upsilon g$ for some $g \in H^2(\mathbb{C}^r)$. Clearly, $H_\Phi f$ is a maximizing vector of H_Φ^*. As we have mentioned in §2, $H_\Phi f = (\Phi - F)f$. Hence, $\bar{z}\overline{\Phi - F}\bar{f}$ is a maximizing vector of $H_{\Phi^{\rm t}}$. Therefore $\bar{z}\overline{\Phi - F}\bar{f} \in \Omega H^2(\mathbb{C}^r)$, and so

$$\Omega^{\rm t}(\Phi - F)f = \Omega^{\rm t}(\Phi - F)\Upsilon g = Ug \in H^2_-(\mathbb{C}^r).$$

It follows that g is a maximizing vector of H_U and $\|H_U\| = 1$. Therefore

$$\Upsilon H^2(\mathbb{C}^r) \subset \Upsilon\mathcal{L}.$$

Hence, $\mathcal{L} = H^2(\mathbb{C}^r)$, and by Theorem 4.1, $t_0(U) = \cdots = t_{r-1}(U) = 1$. It follows that U is very badly approximable. Hence, the zero matrix function is the only best approximation of U by analytic matrix functions.

This uniqueness property together with (4.7) implies that U does not depend on the choice of the best approximation F.

It is evident from (4.6) that $\|\Psi\|_{L^\infty} \le 1$. It remains to prove that $\|H_\Psi\| = t_r(\Phi)$.

Suppose that $F_\$$ is another best approximation of Φ by bounded analytic matrix functions. Then as we have already proved, $\Phi - F_\$$ can be represented as

$$\Phi - F_\$ = \mathcal{W}^* \begin{pmatrix} U & 0 \\ 0 & \Psi_\$ \end{pmatrix} \mathcal{V}^*,$$

where $\Psi_\$$ is a matrix function in $L^\infty(\mathbb{M}_{m-r,n-r})$ such that $\|\Psi_\$\|_{L^\infty} \le 1$. Clearly, $s_j((\Phi - F_\$)(\zeta)) = 1$, $0 \le j \le r-1$, and $s_r((\Phi - F_\$)(\zeta)) = \|\Psi_\$(\zeta)\|$ for almost all $\zeta \in \mathbb{T}$.

By Theorem 3.4, a matrix function $G \in H^\infty(\mathbb{M}_{m,n})$ is a best approximation of Φ if and only if there exists $Q \in H^\infty(\mathbb{M}_{m-r,n-r})$ such that $\|\Psi - Q\|_{L^\infty} \le 1$ and

$$\Phi - G = \mathcal{W}^* \begin{pmatrix} U & 0 \\ 0 & \Psi - Q \end{pmatrix} \mathcal{V}^*.$$

This proves that $\|H_\Psi\| = t_r(\Phi)$. ■

Remark. It can be shown easily that if U is an $r \times r$ unitary-valued very badly approximable matrix function and $\|H_U\|_{\rm e} < 1$, then the Toeplitz operator T_U is Fredholm. Indeed, by Theorem 3.1 of [PY1], if Φ is a very badly approximable function in $(H^\infty + C)(\mathbb{M}_{m,n})$ and $\operatorname{rank} \Phi(\zeta) = m$ almost everywhere on $\mathbb{T}$, then the Toeplitz operator $T_{z\Phi} : H^2(\mathbb{C}^n) \to H^2(\mathbb{C}^m)$ has dense range. The results of [PT] show that the proof given in [PY1] also works in the more general case when $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and $\|H_\Phi\|_{\rm e} < \|H_\Phi\|$. Hence, T_{zU} has dense range. Then the operator $H^*_{\bar z U^*} H_{\bar z U^*}$ is unitarily equivalent to the restriction of $H^*_{zU} H_{zU}$ to the orthogonal complement to the subspace

$$\{f \in H^2(\mathbb{C}^r) : \ \|H_{zU} f\|_2 = \|f\|_2\}$$

(see [Pe2]). Since $\|H_U\|_{\rm e} < 1$, this subspace is finite-dimensional, and so

$$\|H_{U^*}\|_{\rm e} = \|H_{\bar z U^*}\|_{\rm e} = \lim_{j\to\infty} s_j(H_{\bar z U^*}) = \lim_{j\to\infty} s_j(H_{zU}) = \|H_{zU}\|_{\rm e} = \|H_U\|_{\rm e}.$$

The result follows now from the well-known fact that for a unitary-valued function U the conditions $\|H_U\|_{\rm e} < 1$ and $\|H_{U^*}\|_{\rm e} < 1$ are equivalent to the fact that T_U is Fredholm (see e.g., §1 of [AP2] for some comments).

Factorizations of the form (4.6) with Ψ satisfying

$$\|\Psi\|_{L^\infty} \le t_0 \quad \text{and} \quad \|H_\Psi\| < t_0$$

form a special class of *partial canonical factorizations*. The matrix function Ψ is called the *residual entry* of the partial canonical factorization. The notion of a partial canonical factorization in the general case will be defined in §7.

The following theorem together with Theorem 4.3 gives a characterization of the badly approximable matrix functions Φ satisfying the condition $\|H_\Phi\|_{\rm e} < \|H_\Phi\|$.

Theorem 4.6. *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. Suppose that Φ admits a representation of the form*

$$\Phi = \mathcal{W}^* \begin{pmatrix} \sigma U & 0 \\ 0 & \Psi \end{pmatrix} \mathcal{V}^*,$$

where $\sigma > 0$, $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$ are r-balanced matrix functions, U is a very badly approximable unitary-valued $r \times r$ matrix function such that $\|H_U\|_{\mathrm{e}} < 1$, and $\|\Psi\|_{L^\infty} \le \sigma$. Then Φ is badly approximable and $t_0(\Phi) = \cdots = t_{r-1}(\Phi) = \sigma$.

Proof. Suppose that $\mathcal{V}$ and $\mathcal{W}$ are given by (4.5). Let $g \in H^2(\mathbb{C}^r)$ be a maximizing vector of H_U. Then it is easy to see that

$$\|H_\Phi \Upsilon g\|_2 = \sigma \|\Upsilon g\|,$$

while

$$\|H_\Phi\| \le \|\Phi\|_{L^\infty} = \sigma.$$

Hence, $\|H_\Phi\| = \sigma$, Υg is a maximizing vector of H_Φ, and Φ is badly approximable.

Since U is very badly approximable, it follows from Theorem 4.1 that the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^r)$ that contains all maximizing vectors of H_U is the space $H^2(\mathbb{C}^r)$ itself. Let $\mathcal{M}$ be the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^n)$ that contains all maximizing vectors of H_Φ. Since Υ is co-outer, it follows that the matrix $\Upsilon(\zeta)$ has rank r for all $\zeta \in \mathbb{D}$. Hence,

$$\dim\{f(\zeta) : \ f \in \mathcal{M}\} \ge r \quad \text{for all} \quad \zeta \in \mathbb{D}.$$

It follows now from Theorem 4.1 that $t_0(\Phi) = \cdots = t_{r-1}(\Phi)$. ■

We also need the following version of the converse to Theorem 4.3.

Theorem 4.7. *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. Suppose that Φ admits a representation of the form*

$$\Phi = \mathcal{W}^* \begin{pmatrix} \sigma U & 0 \\ 0 & \Psi \end{pmatrix} \mathcal{V}^*,$$

where $\sigma > 0$, $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$ are r-balanced matrix functions of the form (4.5), U is a very badly approximable unitary-valued $r \times r$ matrix function such that $\|H_U\|_{\mathrm{e}} < 1$, and $\|H_\Psi\| < \sigma$. Then $\Upsilon H^2(\mathbb{C}^r)$ is the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^n)$ that contains all maximizing vectors of H_Φ and $\Omega H^2(\mathbb{C}^r)$ is the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^m)$ that contains all maximizing vectors of $H_{\Phi^{\mathrm{t}}}$.

Proof. By Theorem 3.4, we may assume without loss of generality that $\|\Psi\|_{L^\infty} < s$. We need the following lemma.

Lemma 4.8. *Suppose that Φ satisfies the hypotheses of Theorem 4.7. A function f in $H^2(\mathbb{C}^n)$ is a maximizing vector of H_Φ if and only if $f = \Upsilon g$, where $g \in H^2(\mathbb{C}^r)$ and g is a maximizing vector of H_U.*

Let us first complete the proof of Theorem 4.7. Since U is very badly approximable, we have $t_0(U) = \cdots = t_{r-1}(U) = 1$. By Theorem 4.1, the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^r)$ that contains all maximizing vectors of H_U is $H^2(\mathbb{C}^r)$. It follows now from Lemma 4.8 that the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^n)$ that contains all maximizing vectors of H_Φ is $\Upsilon H^2(\mathbb{C}^r)$. To complete the proof, we can apply this result to the matrix function Φ^{t}. ■

Proof of Lemma 4.8. First of all, by Theorem 4.6, $\|H_\Phi\| = \sigma$. Without loss of generality we may assume that $\sigma = 1$. It has been proved in the proof of Theorem 4.6 that if g is a maximizing vector of H_U, then Υg is a maximizing vector of H_Φ.

Suppose now that f is a maximizing vector of H_Φ. We have

$$\begin{aligned} \Phi f &= \mathcal{W}^* \begin{pmatrix} U & 0 \\ 0 & \Psi \end{pmatrix} \mathcal{V}^* f \\ &= \mathcal{W}^* \begin{pmatrix} U & 0 \\ 0 & \Psi \end{pmatrix} \begin{pmatrix} \Upsilon^* f \\ \Theta^{\mathrm{t}} f \end{pmatrix} \\ &= \mathcal{W}^* \begin{pmatrix} U\Upsilon^* f \\ \Psi\Theta^{\mathrm{t}} f \end{pmatrix}. \end{aligned}$$

Since $\mathcal{W}^*$ is unitary-valued and $\|\Psi\|_{L^\infty} < 1$, it follows that $\Theta^{\mathrm{t}} f = 0$. Put $g = \Upsilon^* f \in L^2(\mathbb{C}^r)$. We have

$$f = \mathcal{V}\mathcal{V}^* f = \mathcal{V} \begin{pmatrix} \Upsilon^* \\ \Theta^{\mathrm{t}} \end{pmatrix} f = \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix} \begin{pmatrix} \Upsilon^* f \\ 0 \end{pmatrix} = \Upsilon\Upsilon^* f = \Upsilon g.$$

Let us show that $g \in H^2(\mathbb{C}^r)$. Let γ be a vector in $\mathbb{C}^r$. Since Υ^{t} is outer, it follows that there exists a sequence $\{\varphi_j\}$ of functions in $H^2(\mathbb{C}^n)$ such that $\{\Upsilon^{\mathrm{t}}\varphi_j\}$ converges in $H^2(\mathbb{C}^r)$ to the function identically equal to γ, and so the sequence $\{\varphi_j^{\mathrm{t}}\Upsilon\}$ converges to the function identically equal to γ^{t}. Hence,

$$\lim_{j\to\infty} \{\varphi_j^{\mathrm{t}} f\} = \lim_{j\to\infty} \{\varphi_j^{\mathrm{t}} \Upsilon g\} = \gamma^{\mathrm{t}} g$$

in L^1, and so $\gamma^{\mathrm{t}} g \in H^2$ for any constant vector γ. Consequently, $g \in H^2(\mathbb{C}^r)$.

We have

$$\Phi f = \mathcal{W}^* \begin{pmatrix} U\Upsilon^*\Upsilon g \\ 0 \end{pmatrix} = \begin{pmatrix} \overline{\Omega} & \Xi \end{pmatrix} \begin{pmatrix} Ug \\ 0 \end{pmatrix} = \begin{pmatrix} \overline{\Omega} U g \\ 0 \end{pmatrix}.$$

Clearly, f is a maximizing vector of H_Φ if and only if $\overline{\Omega} U g \in H^2_-(\mathbb{C}^r)$ which is equivalent to the condition $\bar{z}\Omega\overline{Ug} \in H^2(\mathbb{C}^r)$. Since Ω^{t} is outer, we can apply the same reasoning as above to show that $\bar{z}\overline{Ug} \in H^2(\mathbb{C}^r)$ which is equivalent to the fact that $Ug \in H^2_-(\mathbb{C}^r)$. But the latter just means that $g \in \operatorname{Ker} T_U$, and so g is a maximizing vector of H_U. ■

Suppose now that the matrix function Ψ in the factorization (4.6) also satisfies the condition $\|H_\Psi\|_{\mathrm{e}} < \|H_\Psi\|$. Then we can continue this process, find a best

analytic approximation G of Ψ and factorize $\Psi - G$ as in (4.6). If we are able to continue this diagonalization process till the very end, we construct the unique superoptimal approximation Q of Φ and obtain a canonical factorization of $\Phi - Q$.

Therefore we need an estimate of $\|H_\Psi\|_{\mathrm{e}}$. In [PT] in the case $r = 1$ it was shown that $\|H_\Psi\|_{\mathrm{e}} \le \|H_\Phi\|_{\mathrm{e}}$. We want to obtain the same inequality for an arbitrary r. We could try to generalize the proof given in [PT] to the case of an arbitrary r. However, we are going to choose another way. We would like to deduce this result for an arbitrary r from the corresponding result in the case $r = 1$.

5. **Relations with thematic factorizations**

In this section we compare partial canonical factorizations obtained in §4 with partial thematic factorizations and find useful relations between the complementing matrix functions Θ and Ξ in (4.5) and the corresponding complementing matrix functions in partial thematic factorizations.

Let Φ be a matrix function in $L^\infty(\mathbb{M}_{m,n})$ such that $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. Let r be the number of superoptimal singular values of Φ equal to $\|H_\Phi\|$. Suppose that $r < \min\{m, n\}$. It follows from the results of §4 that if F is a best approximation of Φ by bounded analytic matrix functions, then $\Phi - F$ admits a partial canonical factorization

$$\Phi - F = \mathcal{W}^* \begin{pmatrix} t_0 U & 0 \\ 0 & \Psi \end{pmatrix} \mathcal{V}^*, \tag{5.1}$$

where $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$ are r-balanced matrix functions of the form (4.5) and $\|\Psi\|_{L^\infty} \le t_0 = \|H_\Phi\|$, and $\|H_\Psi\| < t_0$.

On the other hand, it follows from the results of [PT] that $\Phi - F$ admits a partial thematic factorization of the form

$$\Phi - F = W_0^* \cdots W_{r-1}^* \begin{pmatrix} t_0 u_0 & 0 & \cdots & 0 & 0 \\ 0 & t_0 u_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & t_0 u_{r-1} & 0 \\ 0 & 0 & \cdots & 0 & \Delta \end{pmatrix} V_{r-1}^* \cdots V_0^*, \tag{5.2}$$

where $\|\Delta\|_{L^\infty} \le t_0$ and $\|H_\Delta\| < t_0$,

$$V_j = \begin{pmatrix} I_j & 0 \\ 0 & \breve{V}_j \end{pmatrix}, \quad W_j = \begin{pmatrix} I_j & 0 \\ 0 & \breve{W}_j \end{pmatrix}, \quad 0 \le j \le r - 1,$$

with thematic matrix functions $\breve{V}_j, \breve{W}_j^{\mathrm{t}}$, and the u_j are unimodular functions such that the Toeplitz operators T_{u_j} are Fredholm with $\operatorname{ind} T_{u_j} > 0$. For $j = 0$ we assume that $\breve{V}_0 = V_0$ and $\breve{W}_0 = W_0$.

Suppose that

$$\breve{V}_j = \begin{pmatrix} \boldsymbol{v}_j & \overline{\Theta}_j \end{pmatrix}, \quad \breve{W}_j^{\mathrm{t}} = \begin{pmatrix} \boldsymbol{w}_j & \Xi_j \end{pmatrix}, \quad 0 \le j \le r-1, \tag{5.3}$$

and

$$\mathcal{V} = \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}, \quad \mathcal{W}^{\mathrm{t}} = \begin{pmatrix} \Omega & \Xi \end{pmatrix},$$

where the matrix functions $\boldsymbol{v}_j, \Theta_j, \boldsymbol{w}_j, \Xi_j, \Upsilon, \Theta, \Omega, \Xi$ are inner and co-outer.

Theorem 5.1. *Under the above hypotheses there exist constant unitary matrices $\mathfrak{U}_1 \in \mathbb{M}_{\mathfrak{n}-\mathfrak{r},\mathfrak{n}-\mathfrak{r}}$ and $\mathfrak{U}_2 \in \mathbb{M}_{\mathfrak{m}-\mathfrak{r},\mathfrak{m}-\mathfrak{r}}$ such that*

$$\Theta = \Theta_0\Theta_1\cdots\Theta_{r-1}\mathfrak{U}_1 \tag{5.4}$$

and

$$\Xi = \Xi_0\Xi_1\cdots\Xi_{r-1}\mathfrak{U}_2. \tag{5.5}$$

Proof. It follows from Theorem 3.4 that if we replace F with another best approximation G, the matrix function $\Phi - G$ will still admit factorizations of the forms (5.1) and (5.2) with the same matrix functions $\mathcal{V}, \mathcal{W}, V_j, W_j$. Hence, we may assume that $F \in \boldsymbol{\Omega}_r$ (see §1). Then the matrix functions Ψ in (5.1) and Δ in (5.2) satisfy the inequalities

$$\|\Psi\|_{L^\infty} < t_0 \quad \text{and} \quad \|\Delta\|_{L^\infty} < t_0.$$

Define the function $\rho : \mathbb{R} \to \mathbb{R}$ by

$$\rho(t) = \begin{cases} t, & t \ge t_0^2 \\ 0, & t < t_0^2 \end{cases}.$$

Consider the operator $M : H^2(\mathbb{C}^n) \to L^2(\mathbb{C}^n)$ of multiplication by the matrix function $\rho\big((\Phi - F)^{\mathrm{t}}(\overline{\Phi - F})\big)$. It was proved in Lemmas 3.4 and 3.5 of [AP1] that

$$\operatorname{Ker} M = \Theta_0\Theta_1\cdots\Theta_{r-1}H^2(\mathbb{C}^{n-r}). \tag{5.6}$$

On the other hand, it follows from (5.1) that

$$(\Phi - F)^{\mathrm{t}}(\overline{\Phi - F}) = \overline{\mathcal{V}} \begin{pmatrix} t_0^2 I_r & 0 \\ 0 & \Psi^{\mathrm{t}}\overline{\Psi} \end{pmatrix} \mathcal{V}^{\mathrm{t}}$$

and since $\|\Psi^{\mathrm{t}}\overline{\Psi}\|_{L^\infty} < t_0^2$, we have

$$\rho\big((\Phi - F)^{\mathrm{t}}(\overline{\Phi - F})\big) = \overline{\mathcal{V}} \begin{pmatrix} t_0^2 I_r & 0 \\ 0 & 0 \end{pmatrix} \mathcal{V}^{\mathrm{t}} = \overline{\mathcal{V}} \begin{pmatrix} t_0^2 I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Upsilon^{\mathrm{t}} \\ \Theta^* \end{pmatrix}.$$

By Theorem 3.1, $\operatorname{Ker} T_{\Upsilon^{\mathrm{t}}} = \Theta H^2(\mathbb{C}^{n-r})$. It is easy to see now that $\operatorname{Ker} M = \Theta H^2(\mathbb{C}^{n-r})$. Together with (5.6) this yields

$$\Theta_0\Theta_1\cdots\Theta_{r-1}H^2(\mathbb{C}^{n-r}) = \Theta H^2(\mathbb{C}^{n-r})$$

which means that both inner functions Θ and $\Theta_0\Theta_1\cdots\Theta_{r-1}$ determine the same invariant subspace of multiplication by z on $H^2(\mathbb{C}^n)$. Therefore there exists a

constant unitary function $\mathfrak{U}_1$ such that (5.4) holds (see §2). To prove (5.5) we can apply (5.4) to $(\Phi - F)^{\mathrm{t}}$. ■

Corollary 5.2. *Let Ψ and Δ be the matrix functions in the factorizations* (5.1) *and* (5.2). *Then*

$$\Delta = \mathfrak{U}_2 \Psi \mathfrak{U}_1^t, \tag{5.7}$$

where $\mathfrak{U}_1$ and $\mathfrak{U}_2$ are unitary matrices from (5.4) *and* (5.5).

Proof. By Corollary 3.2 of [AP1],

$$\Delta = \Xi_{r-1}^* \cdots \Xi_1^* \Xi_0^* (\Phi - F) \overline{\Theta_0 \Theta_1 \cdots \Theta_{r-1}}.$$

By Theorem 5.1,

$$\Delta = \mathfrak{U}_2 \Xi^* (\Phi - F) \overline{\Theta} \mathfrak{U}_1^{\mathrm{t}}.$$

On the other hand, it is easy to see from (5.1) that

$$\Psi = \Xi^* (\Phi - F) \overline{\Theta}, \tag{5.8}$$

which implies (5.7). ■

Now we are in a position to estimate $\|H_\Psi\|_{\mathrm{e}}$ for the residual entry Ψ in the factorization (5.1). This will be used in §7 to obtain canonical factorizations of very badly approximable matrix functions.

Theorem 5.3. *Let Φ be a function in $L^\infty(\mathbb{M}_{m,n})$ such that $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. Then for the residual entry Ψ in the partial canonical factorization* (5.1) *the following inequality holds*

$$\|H_\Psi\|_{\mathrm{e}} \le \|H_\Phi\|_{\mathrm{e}}.$$

Proof. Iterating Theorem 6.3 of [PT], we find that $\|H_\Delta\|_{\mathrm{e}} \le \|H_\Phi\|_{\mathrm{e}}$. The result follows now from (5.7). ■

Consider now the unitary-valued matrix function U in the partial canonical factorization (5.1). As we have observed in the remark following the proof of Theorem 4.3, the Toeplitz operator T_U is Fredholm. We are going to evaluate now the index of T_U in terms of the indices of the partial thematic factorization (5.2). Recall that the indices k_j of the partial thematic factorization (5.2) are defined by

$$k_j \stackrel{\text{def}}{=} \operatorname{ind} T_{u_j}, \quad 0 \le j \le r-1.$$

Theorem 5.4. *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. Then the entry U of of the diagonal block matrix function in the partial canonical factorization* (5.1) *satisfies*

$$\operatorname{ind} T_U = \dim \operatorname{Ker} T_U = k_0 + k_1 + \cdots + k_{r-1}.$$

Proof. Without loss of generality we may assume that $\|H_\Phi\| = 1$. By Theorem 4.3, U is very badly approximable, $\|H_U\|_{\mathrm{e}} < \|H_U\| = 1$. Therefore the Toeplitz operator T_{zU} has dense range in $H^2(\mathbb{C}^r)$ (see Theorem 3.1 of [PY1] where this fact was proved in the case $U \in (H^\infty + C)(\mathbb{M}_{r,r})$, however the proof given in [PY1] also works in the more general case $\|H_U\|_{\mathrm{e}} < \|H_U\| = 1$). Therefore $\operatorname{Ker} T_U^* = \{0\}$, and so $\operatorname{ind} T_U = \dim \operatorname{Ker} T_U$.

By Theorem 9.3 of [PT],

$$\dim\{f \in H^2(\mathbb{C}^n) : \ \|H_\Phi f\|_2 = \|f\|_2\} = k_0 + k_1 + \cdots + k_{r-1} \tag{5.9}$$

(earlier this result was proved in [PY2], Theorem 2.2 in the case $\Phi \in (H^\infty + C)(\mathbb{M}_{m,n})$). Let us show that the left-hand side of (5.9) is equal to $\dim \operatorname{Ker} T_U$. Indeed, if $g \in H^2(\mathbb{C}^r)$, then $g \in \operatorname{Ker} T_U$ if and only if g is a maximizing vector of H_U. By Lemma 4.8,

$$\dim\{g \in H^2(\mathbb{C}^r) : \ \|H_U g\|_2 = \|g\|_2\} = \dim\{f \in H^2(\mathbb{C}^n) : \ \|H_\Phi f\|_2 = \|f\|_2\}$$

which proves the result. ■

To conclude this section, we obtain an inequality between the singular values of H_Φ and the singular values of H_Ψ in the factorization (5.1).

Theorem 5.5. *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. Suppose that $F \in H^\infty(\mathbb{M}_{m,n})$ is a best approximation of Φ by bounded analytic functions and $\Phi - F$ is represented by the partial canonical factorization* (5.1). *Then*

$$s_j(H_\Psi) \le s_{\iota+j}(H_\Phi), \quad j \ge 0, \tag{5.10}$$

where $\iota \stackrel{\text{def}}{=} \operatorname{ind} T_U$.

Note that (5.10) was proved in [PT], Theorem 10.1 in the case $r = 1$. We reduce the general case to the case $r = 1$.

Proof. Consider the partial thematic factorization (5.2). Let $k_j \stackrel{\text{def}}{=} \operatorname{ind} T_{u_j}$, $0 \le j \le r-1$, be the indices of this factorization. We have

$$\Phi - F = W_0^* \begin{pmatrix} t_0 u_0 & 0 \\ 0 & \Phi^{[1]} \end{pmatrix} V_0^*, \tag{5.11}$$

where $\Phi^{[1]}$ is given by the partial thematic factorization

$$\Phi^{[1]} = \breve{W}_1^* \cdots \begin{pmatrix} I_{r-2} & 0 \\ 0 & \breve{W}_{r-1}^* \end{pmatrix} \begin{pmatrix} t_0 u_1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & t_0 u_{r-1} & 0 \\ 0 & \cdots & 0 & \Delta \end{pmatrix} \begin{pmatrix} I_{r-2} & 0 \\ 0 & \breve{V}_{r-1}^* \end{pmatrix} \cdots \breve{V}_1^*.$$

Now we can apply Theorem 10.1 of [PT] to the factorization (5.11) and find that

$$s_j(H_{\Phi^{[1]}}) \le s_{k_0+j}(H_\Phi), \quad j \ge 0.$$

Then we can apply Theorem 10.1 of [PT] to the above partial thematic factorization of $\Phi^{[1]}$, *etc.* After applying Theorem 10.1 of [PT] r times we obtain the inequality

$$s_j(H_\Delta) \le s_{k_0+\cdots+k_{r-1}+j}(H_\Phi), \quad j \ge 0.$$

The result follows now from Corollary 5.2 and Theorem 5.4. ■

6. Invertibility of the Toeplitz operators $T_{\mathcal{V}}$ and $T_{\mathcal{W}}$

For a matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$ satisfying $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$ we have constructed in §4 balanced matrix functions $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$. We have mentioned in §3 that by Lemma 6.2 of [Pe4], $T_{\mathcal{V}}$ and $T_{\mathcal{W}}$ have trivial kernels and dense ranges. In this section we prove that under the condition $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$ the Toeplitz operators $T_{\mathcal{V}}$ and $T_{\mathcal{W}}$ are invertible. In the case $r = 1$ it was proved in [PT] (see the proof of Theorem 5.1 of [PT]) that the Toeplitz operators $T_{\mathcal{V}}$ and $T_{\mathcal{W}^{\mathrm{t}}}$ are invertible. Instead of trying to generalize the proof given in [PT] to the case of an arbitrary r we are going to reduce the general case to the case $r = 1$.

Then we are going to prove that the matrix functions $\Upsilon, \Theta, \Omega, \Xi$ given by (4.5) are left invertible in H^∞. Note that this result in the case $r = 1$ plays an important role (see [PY2] and [PT]).

We are going to use the following fact from [Pe2] (see also [PK] where the scalar case was considered):

If V is an $n \times n$ unitary-valued function such that the Toeplitz operator $T_V : H^2(\mathbb{C}^n) \to H^2(\mathbb{C}^n)$ has trivial kernel and dense range, then the operators $H^*_{V^*}H_{V^*}$ and $H^*_V H_V$ are unitarily equivalent. In particular, $\|H_V\| = \|H_{V^*}\|$.

We need another well known fact (see [D] or [Ni] for the scalar case, in the matrix case the proof is the same):

If V is an $n \times n$ unitary-valued function, then T_V is invertible on $H^2(\mathbb{C}^n)$ if and only if $\|H_V\| < 1$ and $\|H_{V^*}\| < 1$.

Consider now a balanced matrix function $\mathcal{V} = \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}$. Clearly, $\|H_{\mathcal{V}}\| = \|H_{\overline{\Theta}}\|$ and $\|H_{\mathcal{V}^*}\| = \|H_{\overline{\Upsilon}}\|$. Since $T_{\mathcal{V}}$ has trivial kernel and dense range (Lemma 6.2 of [Pe4]), it follows that

$$\|H_{\overline{\Theta}}\| = \|H_{\overline{\Upsilon}}\|. \tag{6.1}$$

$T_{\mathcal{V}}$ is invertible if and only if the norms in (6.1) are less than 1. It is also clear that the last condition is equivalent to the invertibility of $T_{\mathcal{V}^{\mathrm{t}}}$.

Theorem 6.1. *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and suppose that $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. Let $\mathcal{V}$ and $\mathcal{W}^t$ be the r-balanced matrix functions in the partial canonical factorization* (5.1). *Then the Toeplitz operators $T_{\mathcal{V}}$, $T_{\mathcal{V}^{\mathrm{t}}}$, $T_{\mathcal{W}}$, and $T_{\mathcal{W}^{\mathrm{t}}}$ are invertible.*

Recall that for $r = 1$ this was proved in [PT] (see the proof of Theorem 5.1 of [PT]). We are going to reduce the general case to the case $r = 1$.

Proof. Consider the partial thematic factorization (5.2) and the corresponding thematic matrix functions $\breve{V}_j$ and the inner matrix functions Θ_j given by (5.3). As we have already mentioned, it was shown in the proof of Theorem 5.1 of [PT] that the Toeplitz operators $T_{\breve{V}_j}$ are invertible for $0 \le j \le r-1$. Since the $\breve{V}_j$ are 1-balanced, this is equivalent to the fact that $\|H_{\overline{\Theta}_j}\| < 1$, $0 \le j \le r-1$ (see the discussion preceding the statement of Theorem 6.1). To prove that $T_{\mathcal{V}}$ is invertible, it is sufficient to show that $\|H_{\overline{\Theta}}\| < 1$, where as usual, $\mathcal{V} = \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}$.

By Theorem 5.1, $\Theta = \Theta_0 \cdots \Theta_{r-1}\mathfrak{U}$, where $\mathfrak{U}$ is a constant unitary matrix. Clearly, to show that $\|H_{\overline{\Theta}}\| < 1$, it is sufficient to prove the following lemma.

Lemma 6.2. *Let Θ_1 be an $n \times k$ inner matrix function and let Θ_2 be a $k \times l$ inner matrix function such that $\|H_{\overline{\Theta}_1}\| < 1$ and $\|H_{\overline{\Theta}_2}\| < 1$. Let $\Theta = \Theta_1\Theta_2$. Then $\|H_{\overline{\Theta}}\| < 1$.*

Let us first complete the proof of Theorem 6.1. Applying Lemma 6.2 inductively, we find that $\|H_{\overline{\Theta}}\| < 1$. As we have explained this is equivalent to the invertibility of $T_{\mathcal{V}}$ and $T_{\mathcal{V}^{\mathrm{t}}}$. Similarly, one can prove that the Toeplitz operators $T_{\mathcal{W}^{\mathrm{t}}}$ and $T_{\mathcal{W}}$ are invertible. ■

Proof of Lemma 6.2. Let $\sigma < 1$ be a positive number such that $\|H_{\overline{\Theta}_1}\| < \sigma$ and $\|H_{\overline{\Theta}_2}\| < \sigma$. Let $f \in H^2(\mathbb{C}^l)$. We have

$$\begin{aligned} \|H_{\overline{\Theta}}f\|_2^2 &= \|\mathbb{P}_-\overline{\Theta_1\Theta_2}f\|_2^2 \\ &= \|\mathbb{P}_-\overline{\Theta_1}\mathbb{P}_-\overline{\Theta_2}f + \mathbb{P}_-\overline{\Theta_1}\mathbb{P}_+\overline{\Theta_2}f\|_2^2 \\ &= \|\overline{\Theta_1}\mathbb{P}_-\overline{\Theta_2}f + \mathbb{P}_-\overline{\Theta_1}\mathbb{P}_+\overline{\Theta_2}f\|_2^2 \\ &= \|\mathbb{P}_-\overline{\Theta_2}f + \Theta_1^{\mathrm{t}}\mathbb{P}_-\overline{\Theta_1}\mathbb{P}_+\overline{\Theta_2}f\|_2^2. \end{aligned}$$

We claim that the functions $\mathbb{P}_-\overline{\Theta_2}f$ and $\Theta_1^{\mathrm{t}}\mathbb{P}_-\overline{\Theta_1}\mathbb{P}_+\overline{\Theta_2}f$ are orthogonal. Indeed, let $\varphi \in H^2_-(\mathbb{C}^k)$ and $\psi \in H^2(\mathbb{C}^k)$. We have

$$(\varphi, \Theta_1^{\mathrm{t}}\mathbb{P}_-\overline{\Theta_1}\psi) = (\overline{\Theta_1}\varphi, \mathbb{P}_-\overline{\Theta_1}\psi) = (\overline{\Theta_1}\varphi, \overline{\Theta_1}\psi) = (\varphi, \psi) = 0,$$

since $\overline{\Theta_1}$ takes isometric values almost everywhere on $\mathbb{T}$. It follows that

$$\begin{aligned} \|H_{\overline{\Theta}}f\|_2^2 &= \|\mathbb{P}_-\overline{\Theta_2}f\|_2^2 + \|\Theta_1^{\mathrm{t}}\mathbb{P}_-\overline{\Theta_1}\mathbb{P}_+\overline{\Theta_2}f\|_2^2 \\ &\le \|\mathbb{P}_-\overline{\Theta_2}f\|_2^2 + \|\mathbb{P}_-\overline{\Theta_1}\mathbb{P}_+\overline{\Theta_2}f\|_2^2 \\ &= \|\mathbb{P}_-\overline{\Theta}_2f\|_2^2 + \|H_{\overline{\Theta_1}}\mathbb{P}_+\overline{\Theta}_2f\|_2^2 \\ &\le \|\mathbb{P}_-\overline{\Theta}_2f\|_2^2 + \sigma^2\|\mathbb{P}_+\overline{\Theta}_2f\|_2^2 \\ &= \sigma^2(\|\mathbb{P}_-\overline{\Theta}_2f\|_2^2 + \|\mathbb{P}_+\overline{\Theta}_2f\|_2^2) + (1-\sigma^2)\|\mathbb{P}_-\overline{\Theta}_2f\|_2^2 \\ &= \sigma^2\|\overline{\Theta_2}f\|_2^2 + (1-\sigma^2)\|H_{\overline{\Theta}_2}f\|_2^2 \\ &\le \sigma^2\|f\|_2^2 + \sigma^2(1-\sigma^2)\|f\|_2^2 = (2\sigma^2 - \sigma^4)\|f\|_2^2. \end{aligned}$$

The result follows now from the trivial inequality $2\sigma^2 - \sigma^4 < 1$. ■

Now we are going to prove that the matrix functions $\Upsilon, \Theta, \Omega, \Xi$ are left invertible in H^∞. Recall that a matrix function $\Delta \in H^\infty(\mathbb{M}_{\iota,\kappa})$ is called *left invertible in H^∞* if there exists a matrix function $\Lambda \in H^\infty(\mathbb{M}_{\kappa,\iota})$ such that $\Lambda\Delta$ is identically equal to I_κ. Recall that by a theorem of Sz.-Nagy and Foias [SNF2], Δ is left invertible in H^∞ if and only if the Toeplitz operator $T_{\overline{\Delta}} : H^2(\mathbb{C}^\kappa) \to H^2(\mathbb{C}^\iota)$ is bounded below, i.e., $\|T_{\overline{\Delta}}\varphi\|_2 \ge \delta\|\varphi\|_2$ for some $\delta > 0$.

Theorem 6.3. *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. Suppose that $\mathcal{V}$ and $\mathcal{W}$ are matrix functions of the form* (4.5) *in the partial canonical factorization* (5.1). *Then $\Upsilon, \Theta, \Omega, \Xi$ are left invertible in H^∞.*

Proof. Let us prove that Θ is left invertible. By Theorem 6.1. the Toeplitz operator $T_{\mathcal{V}}$ is invertible. Hence, $\|T_{\mathcal{V}}f\|_2 \ge \delta\|f\|_2$, $f \in H^2(\mathbb{C}^n)$, for some $\delta > 0$. If we apply this inequality to functions f of the form $f = \begin{pmatrix} 0 \\ \varphi \end{pmatrix}$, we find that $\|T_{\overline{\Theta}}\varphi\|_2 \ge \delta\|\varphi\|_2$, $\varphi \in H^2(\mathbb{C}^{n-r})$. It follows now from the Sz.-Nagy–Foias theorem mentioned above that Θ is left invertible in H^∞.

To prove that Υ is left invertible in H^∞, we can apply the above reasoning to the matrix function $\overline{\mathcal{V}}$. Finally, to obtain the same results for Ω and Ξ, we can apply the above reasoning to the matrix functions $\mathcal{W}^*$ and $\mathcal{W}^{\mathrm{t}}$. ■

7. Canonical factorizations of very badly approximable functions

Given $\Phi \in L^\infty(\mathbb{M}_{m,n})$, consider the sequence $\{t_j\}$ of its superoptimal singular values. Suppose that

$$t_0 = \cdots = t_{r_1-1} > t_{r_1} = \cdots = t_{r_2-1} > \cdots > t_{r_{\iota-1}} = \cdots = t_{r_\iota-1} \tag{7.1}$$

are all nonzero superoptimal singular values of Φ, i.e., t_0 has multiplicity r_1 and t_{r_j} has multiplicity $r_{j+1} - r_j$, $1 \le j \le \iota - 1$.

By Theorem 4.3, if $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$ and F_0 is a best approximation of Φ by bounded analytic matrix function, then $\Phi - F_0$ admits a factorization

$$\Phi - F_0 = \mathcal{W}_0^* \begin{pmatrix} t_0 U_0 & 0 \\ 0 & \Phi^{[1]} \end{pmatrix} \mathcal{V}_0^*,$$

where $\mathcal{V}_0$ and $\mathcal{W}_0^{\mathrm{t}}$ are r_1-balanced matrix functions, U_0 is very badly approximable unitary-valued matrix function such that $\|H_{U_0}\|_{\mathrm{e}} < 1$, and $\Phi^{[1]}$ is a matrix function in $L^\infty(\mathbb{M}_{m-r_1,n-r_1})$ such that $\|H_{\Phi^{[1]}}\| = t_{r_1}$.

By Theorem 5.3, $\|H_{\Phi^{[1]}}\|_{\mathrm{e}} \le \|H_\Phi\|_{\mathrm{e}}$. If t_{r_1} is still greater than $\|H_\Phi\|_{\mathrm{e}}$, we can apply Theorem 4.3 to $\Phi^{[1]}$ and find that for a best approximation G_1 of $\Phi^{[1]}$ the

matrix function $\Phi^{[1]} - G_1$ admits a factorization

$$\Phi^{[1]} - G_1 = \breve{\mathcal{W}}_1^* \begin{pmatrix} t_{r_1} U_1 & 0 \\ 0 & \Phi^{[2]} \end{pmatrix} \breve{\mathcal{V}}_1^*,$$

where $\breve{\mathcal{V}}_1$ and $\breve{\mathcal{W}}_1^{\rm t}$ are $(r_2 - r_1)$-balanced matrix functions, U_1 is a very badly approximable unitary-valued matrix function of size $(r_2 - r_1) \times (r_2 - r_1)$ such that $\|H_{U_1}\|_{\rm e} < 1$, and $\Phi^{[2]}$ is a matrix function in $L^\infty(\mathbb{M}_{m-r_2,n-r_2})$ such that $\|H_{\Phi^{[2]}}\| = t_{r_2}$.

We can apply now Theorem 3.4 and find a matrix function $F_1 \in H^\infty(\mathbb{M}_{m,n})$ such that

$$\Phi - F_1 = \mathcal{W}_0^* \mathcal{W}_1^* \begin{pmatrix} t_0 U_0 & 0 & 0 \\ 0 & t_{r_1} U_1 & 0 \\ 0 & 0 & \Phi^{[2]} \end{pmatrix} \mathcal{V}_1^* \mathcal{V}_0^*,$$

where

$$\mathcal{V}_1 = \begin{pmatrix} I_{r_1} & 0 \\ 0 & \breve{\mathcal{V}}_1 \end{pmatrix} \quad \text{and} \quad \mathcal{W}_1 = \begin{pmatrix} I_{r_1} & 0 \\ 0 & \breve{\mathcal{W}}_1 \end{pmatrix}.$$

If t_{r_2} is still greater than $\|H_\Phi\|_{\rm e}$, we can continue this process and apply Theorem 4.3 to $\Phi^{[2]}$. Suppose now that $t_{r_{d-1}} > \|H_\Phi\|_{\rm e}$, $2 \le d \le \iota$. Then continuing the above process and applying Theorem 3.4, we can find a function $F \in H^\infty(\mathbb{M}_{m,n})$ such that $\Phi - F$ admits a factorization

$$\Phi - F = \mathcal{W}_0^* \cdots \mathcal{W}_{d-1}^* \begin{pmatrix} t_0 U_0 & 0 & \cdots & 0 & 0 \\ 0 & t_{r_1} U_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & t_{r_{d-1}} U_{d-1} & 0 \\ 0 & 0 & \cdots & 0 & \Phi^{[d]} \end{pmatrix} \mathcal{V}_{d-1}^* \cdots \mathcal{V}_0^*, \tag{7.2}$$

where the U_j are $(r_{j+1} - r_j) \times (r_{j+1} - r_j)$ very badly approximable unitary valued functions such that $\|H_{U_j}\|_{\rm e} < 1$,

$$\mathcal{V}_j = \begin{pmatrix} I_{r_j} & 0 \\ 0 & \breve{\mathcal{V}}_j \end{pmatrix} \quad \text{and} \quad \mathcal{W}_j = \begin{pmatrix} I_{r_j} & 0 \\ 0 & \breve{\mathcal{W}}_j \end{pmatrix}, \quad 1 \le j \le d-1, \tag{7.3}$$

$\breve{\mathcal{V}}_j$ and $\breve{\mathcal{W}}_j^{\rm t}$ are $(r_{j+1} - r_j)$-balanced matrix functions, and Φ is a matrix function satisfying

$$\|\Phi^{[d]}\|_{L^\infty} \le t_{r_{d-1}}, \quad \text{and} \quad \|H_{\Phi^{[d]}}\| < t_{r_{d-1}}. \tag{7.4}$$

Factorizations of the form (7.2) with $\mathcal{V}_j$ and $\mathcal{W}_j$ of the form (7.3) and $\Phi^{[d]}$ satisfying (7.4) are called *partial canonical factorizations.* The matrix function $\Phi^{[d]}$ is called the *residual entry* of the partial canonical factorization (7.2).

Finally, if $\|H_\Phi\|_{\mathrm{e}}$ is less than the smallest nonzero superoptimal singular value $t_{r_\iota - 1}$, then we can complete this process and construct the unique superoptimal approximation of Φ by bounded analytic matrix functions. This proves the following theorem.

Theorem 7.1. *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and suppose that the nonzero superoptimal singular values of Φ satisfy* (7.1). *If $\|H_\Phi\|_{\mathrm{e}} < t_{r_\iota - 1}$ and F is the unique superoptimal approximation of Φ by bounded analytic functions, then $\Phi - F$ admits a factorization*

$$\Phi - F = \mathcal{W}_0^* \cdots \mathcal{W}_{\iota-1}^* \begin{pmatrix} t_0 U_0 & 0 & \cdots & 0 & 0 \\ 0 & t_{r_1} U_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & t_{r_{\iota-1}} U_{\iota-1} & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \mathcal{V}_{\iota-1}^* \cdots \mathcal{V}_0^*, \tag{7.5}$$

where the $\mathcal{V}_j$ and $\mathcal{W}_j$, and U_j are as above.

Note that the lower right entry of the diagonal matrix function on the right-hand side of (7.5) is the zero matrix function of size $(m - r_\iota) \times (n - r_\iota)$. Here it may happen that $m - r_\iota$ or $n - r_\iota$ can be zero.

Clearly, the left-hand side of (7.5) is a very badly approximable matrix function. Factorizations of the form (7.5) are called *canonical factorizations* of very badly approximable matrix functions.

The following result shows that the right-hand side of (7.5) is always a very badly approximable matrix function.

Theorem 7.2. *Let Φ be a matrix function in $L^\infty(\mathbb{M}_{m,n})$ such that $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$, ι is a positive integer, $r_1, \cdots, r_\iota$ are positive integers satisfying*

$$r_1 < r_2 < \cdots < r_\iota$$

and

$$\sigma_0 > \sigma_1 > \cdots > \sigma_{\iota-1} > 0.$$

Suppose that Φ admits a factorization

$$\Phi = \mathcal{W}_0^* \cdots \mathcal{W}_{\iota-1}^* \begin{pmatrix} \sigma_0 U_0 & 0 & \cdots & 0 & 0 \\ 0 & \sigma_1 U_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_{\iota-1} U_{\iota-1} & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \mathcal{V}_{\iota-1}^* \cdots \mathcal{V}_0^*,$$

in which the U_j, $\mathcal{V}_j$, and $\mathcal{W}_j$ are as above. Then Φ is very badly approximable and the superoptimal singular values of Φ are given by

$$t_\kappa(\Phi) = \begin{cases} \sigma_0, & \kappa < r_1 \\ \sigma_j, & r_j \le \kappa < r_{j+1} \\ 0, & \kappa \ge r_\iota \end{cases}.$$

Proof. Let

$$\mathcal{V}_0 = \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix} \quad \text{and} \quad \mathcal{W}_0 = \begin{pmatrix} \Omega & \Xi \end{pmatrix},$$

where $\Upsilon, \Theta, \Omega, \Xi$ are inner and co-outer matrix functions.

By Theorem 4.7, the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^n)$ that contains all maximizing vectors of H_Φ is $\Upsilon H^2(\mathbb{C}^{r_1})$ and the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^m)$ that contains all maximizing vectors of $H_{\Phi^{\mathrm{t}}}$ is $\Omega H^2(\mathbb{C}^{r_1})$. By Theorem 4.6, Φ is badly approximable and

$$t_0(\Phi) = \cdots = t_{r_1 - 1} = \sigma_0.$$

By Theorem 3.4, we can reduce our problem to the function

$$\breve{\mathcal{W}}_1^* \cdots \begin{pmatrix} I_{r_{\iota-1}-r_1} & 0 \\ 0 & \breve{\mathcal{W}}_{\iota-1}^* \end{pmatrix} \begin{pmatrix} \sigma_1 U_1 & \cdots & \mathbb{O} & \mathbb{O} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbb{O} & \cdots & \sigma_{\iota-1} U_{\iota-1} & \mathbb{O} \\ \mathbb{O} & \cdots & \mathbb{O} & \mathbb{O} \end{pmatrix} \begin{pmatrix} I_{r_{\iota-1}-r_1} & 0 \\ 0 & \breve{\mathcal{V}}_{\iota-1}^* \end{pmatrix} \cdots \breve{\mathcal{V}}_1^*.$$

This function is also represented by a canonical factorization which makes it possible to continue this process and prove that Φ is very badly approximable and the superoptimal singular values of Φ satisfy the desired equality. ■

8. **Invariance properties of canonical factorizations**

In this section we demonstrate an advantage of canonical factorizations over thematic factorizations. Namely, we show that canonical factorizations possess certain invariance properties, i.e., the matrix functions U_j in the canonical factorization (7.5) are uniquely determined modulo unitary constant factors. Moreover, if

$$\mathcal{V}_0 = \begin{pmatrix} \Upsilon_0 & \overline{\Theta}_0 \end{pmatrix}, \quad \mathcal{W}_0^{\mathrm{t}} = \begin{pmatrix} \Omega_0 & \Xi_0 \end{pmatrix},$$

$$\breve{\mathcal{V}}_j = \begin{pmatrix} \Upsilon_j & \overline{\Theta}_j \end{pmatrix}, \quad \breve{\mathcal{W}}_j^{\mathrm{t}} = \begin{pmatrix} \Omega_j & \Xi_j \end{pmatrix}, \quad 1 \le j \le \iota - 1,$$

and the $\breve{\mathcal{V}}_j$ and $\breve{\mathcal{W}}_j$ are given by (7.3), then the matrix functions $\Upsilon_j, \Theta_j, \Omega_j, \Xi_j$ are also uniquely determined modulo constant unitary factors.

We start with partial canonical factorizations of the form (5.1). Suppose that a matrix function Φ in $L^\infty(\mathbb{M}_{m,n})$ satisfies $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$ and admits partial

canonical factorizations

$$\Phi = \begin{pmatrix} \overline{\Omega} & \Xi \end{pmatrix} \begin{pmatrix} \sigma U & 0 \\ 0 & \Psi \end{pmatrix} \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}^* \tag{8.1}$$

and

$$\Phi = \begin{pmatrix} \overline{\Omega^\circ} & \Xi^\circ \end{pmatrix} \begin{pmatrix} \sigma^\circ U^\circ & 0 \\ 0 & \Psi^\circ \end{pmatrix} \begin{pmatrix} \Upsilon^\circ & \overline{\Theta^\circ} \end{pmatrix}^*, \tag{8.2}$$

where $\|\Psi\|_{L^\infty} \le \sigma$, $\|H_\Psi\| < \sigma$, $\|\Psi^\circ\|_{L^\infty} \le \sigma^\circ$, $\|H_{\Psi^\circ}\| < \sigma^\circ$, U is an $r \times r$ very badly approximable unitary-valued function such that $\|H_U\|_{\mathrm{e}} < 1$, U° is an $r^\circ \times r^\circ$ very badly approximable unitary-valued function such that $\|H_{U^\circ}\|_{\mathrm{e}} < 1$, $\begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}$ and $\begin{pmatrix} \Omega & \overline{\Xi} \end{pmatrix}^{\mathrm{t}}$ are r-balanced matrix functions, and $\begin{pmatrix} \Upsilon^\circ & \overline{\Theta^\circ} \end{pmatrix}$ and $\begin{pmatrix} \Omega^\circ & \overline{\Xi^\circ} \end{pmatrix}^{\mathrm{t}}$ are r°-balanced matrix functions.

Theorem 8.1. *Let Φ be a badly approximable function in $L^\infty(\mathbb{M}_{m,n})$ such that $\|H_\Phi\|_{\mathrm{e}} < \|H_\Phi\|$. Suppose that Φ admits factorizations* (8.1) *and* (8.2). *Then $r = r^\circ$, $\sigma = \sigma^\circ$ and there exist unitary matrices $\mathfrak{V}^\#, \mathfrak{V}^\flat \in \mathbb{M}_{\mathfrak{r},\mathfrak{r}}$, $\mathfrak{U}^\# \in \mathbb{M}_{\mathfrak{n}-\mathfrak{r},\mathfrak{n}-\mathfrak{r}}$, $\mathfrak{U}^\flat \in \mathbb{M}_{\mathfrak{m}-\mathfrak{r},\mathfrak{m}-\mathfrak{r}}$ such that*

$$\Upsilon^\circ = \Upsilon \mathfrak{V}^\#, \quad \Omega^\circ = \Omega \mathfrak{V}^\flat, \tag{8.3}$$

$$\Theta^\circ = \Theta \mathfrak{U}^\#, \quad \Xi^\circ = \Xi \mathfrak{U}^\flat, \tag{8.4}$$

and

$$U^\circ = (\mathfrak{V}^\flat)^{\mathrm{t}} \mathfrak{U} \mathfrak{V}^\#, \quad \Psi^\circ = (\mathfrak{U}^\flat)^* \Psi \overline{\mathfrak{U}^\#}. \tag{8.5}$$

Proof. By Theorem 4.6, $\sigma = \sigma^\circ = \|H_\Phi\|$. Next, by Theorem 4.7, the minimal invariant subspace of multiplication by z on $H^2(\mathbb{C}^n)$ that contains all maximizing vectors of H_Φ is equal to $\Upsilon H^2(\mathbb{C}^r)$ and at the same time it is equal to $\Upsilon^\circ H^2(\mathbb{C}^{r^\circ})$. It follows that $r = r^\circ$ and there exists a unitary matrix $\mathfrak{V}^\# \in \mathbb{M}_{\mathfrak{r},\mathfrak{r}}$ such that $\Upsilon^\circ = \Upsilon \mathfrak{V}^\#$. Applying the same reasoning to Φ^{t}, we find a unitary matrix function $\mathfrak{V}^\flat \in \mathbb{M}_{\mathfrak{r},\mathfrak{r}}$ such that $\Omega^\circ = \Omega \mathfrak{V}^\flat$ which proves (8.3).

By Theorem 3.1,

$$\Theta H^2(\mathbb{C}^{n-r}) = \operatorname{Ker} T_{\Upsilon^{\mathrm{t}}} \quad \text{and} \quad \Theta^\circ H^2(\mathbb{C}^{n-r}) = \operatorname{Ker} T_{(\Upsilon^\circ)^{\mathrm{t}}}.$$

By (8.3), $\operatorname{Ker} T_{\Upsilon^{\mathrm{t}}} = \operatorname{Ker} T_{(\Upsilon^\circ)^{\mathrm{t}}}$ which implies that there exist a unitary matrix $\mathfrak{U}^\# \in \mathbb{M}_{\mathfrak{n}-\mathfrak{r},\mathfrak{n}-\mathfrak{r}}$ such that $\Theta^\circ = \Theta \mathfrak{U}^\#$. Applying the same reasoning to Φ^{t}, we find a unitary matrix $\mathfrak{U}^\flat \in \mathbb{M}_{\mathfrak{m}-\mathfrak{r},\mathfrak{m}-\mathfrak{r}}$ such that $\Xi^\circ = \Xi \mathfrak{U}^\flat$ which proves (8.4).

By (4.7),

$$\sigma U = \Omega^{\mathrm{t}} \Phi \Upsilon \quad \text{and} \quad \sigma U^\circ = (\Omega^\circ)^{\mathrm{t}} \Phi \Upsilon^\circ.$$

This implies the first equality in (8.5).

Finally, by (5.8),

$$\Psi = \Xi^* \Phi \overline{\Theta} \quad \text{and} \quad \Psi^\circ = \Xi^{\circ *} \Phi \overline{\Theta^\circ}$$

which completes the proof of (8.5). ■

We can obtain similar results for arbitrary partial canonical factorizations and canonical factorizations. Let us consider in detail the following special case. Suppose that the matrix functions Ψ and Ψ° in (8.1) and (8.2) admit the following partial canonical factorizations

$$\Psi = \begin{pmatrix} \overline{\Omega}_1 & \Xi_1 \end{pmatrix} \begin{pmatrix} \sigma_1 U_1 & 0 \\ 0 & \Delta \end{pmatrix} \begin{pmatrix} \Upsilon_1 & \overline{\Theta}_1 \end{pmatrix}^* \tag{8.6}$$

and

$$\Psi^\circ = \begin{pmatrix} \overline{\Omega_1^\circ} & \Xi_1^\circ \end{pmatrix} \begin{pmatrix} \sigma_1^\circ U_1^\circ & 0 \\ 0 & \Delta^\circ \end{pmatrix} \begin{pmatrix} \Upsilon_1^\circ & \overline{\Theta_1^\circ} \end{pmatrix}^*, \tag{8.7}$$

where $\|\Delta\|_{L^\infty} \le \sigma_1$, $\|H_\Delta\| < \sigma_1$ $\underline{\text{and}}$ $\|\Delta^\circ\|_{L^\infty} < \sigma_1^\circ$, $\|H_{\Delta^\circ}\| < \sigma_1^\circ$.

By Theorem 8.1, $\Psi^\circ = (\mathfrak{U}^\flat)^* \Psi \overline{\mathfrak{U}^\#}$. It follows now from (8.7) that

$$\Psi = \begin{pmatrix} \overline{\overline{\mathfrak{U}^\flat}\Omega_1^\circ} & \mathfrak{U}^\flat \Xi_{\mathfrak{1}}^\circ \end{pmatrix} \begin{pmatrix} \sigma_1^\circ U_1^\circ & 0 \\ 0 & \Delta^\circ \end{pmatrix} \begin{pmatrix} \overline{\mathfrak{U}^\#}\Upsilon_1^\circ & \overline{\mathfrak{U}^\#\Theta_{\mathfrak{1}}^\circ} \end{pmatrix}^* \tag{8.8}$$

which is another partial canonical factorization of Ψ.

We can compare now the factorizations (8.6) and (8.8). By Theorem 8.1, $\sigma_1^\circ = \sigma_1$ and there exist unitary matrices $\mathfrak{V}_{\mathfrak{1}}^\#, \mathfrak{V}_{\mathfrak{1}}^\flat, \mathfrak{U}_{\mathfrak{1}}^\#, \mathfrak{U}_{\mathfrak{1}}^\flat$ such that

$$\Upsilon_1^\circ = (\mathfrak{U}^\#)^{\mathrm{t}} \Upsilon_{\mathfrak{1}} \mathfrak{V}_{\mathfrak{1}}^\#, \quad \Omega_{\mathfrak{1}}^\circ = (\mathfrak{U}^\flat)^{\mathrm{t}} \Omega_{\mathfrak{1}} \mathfrak{V}_{\mathfrak{1}}^\flat,$$

$$\Theta_1^\circ = (\mathfrak{U}^\#)^* \Theta_{\mathfrak{1}} \mathfrak{U}_{\mathfrak{1}}^\#, \quad \Xi_{\mathfrak{1}}^\circ = (\mathfrak{U}^\flat)^* \Xi_{\mathfrak{1}} \mathfrak{U}_{\mathfrak{1}}^\flat,$$

and

$$U_1^\circ = (\mathfrak{V}_{\mathfrak{1}}^\flat)^{\mathrm{t}} \mathfrak{U}_{\mathfrak{1}} \mathfrak{V}_{\mathfrak{1}}^\#, \quad \Delta^\circ = (\mathfrak{U}_{\mathfrak{1}}^\flat)^* \Delta \overline{\mathfrak{U}_{\mathfrak{1}}}^\#.$$

It is easy to see that the same results hold in the case of arbitrary partial canonical factorizations as well as arbitrary canonical factorizations.

9. **Hereditary properties**

In this section we consider the following heredity problem. Suppose that the initial matrix function Φ belongs to a certain function space X. We study the problem of whether all matrix functions in a (partial) canonical factorization of Φ belong to the same space X (by this we certainly mean that all their entries are in X). This is not true for an arbitrary function space X. In particular, it can be shown that this is not true if X is the space $C(\mathbb{T})$ of continuous function. Nevertheless, we prove that this is true for two natural classes of function spaces. The first class consists of so-called $\mathcal{R}$-spaces, i.e., spaces that can be described in terms of rational approximation in the norm of BMO (see [PK]). The second class of spaces consists of Banach algebras satisfying Axioms (A1)–(A4) below. In both

cases we apply so-called recovery theorems for unitary-valued functions obtained in [Pe2] and [Pe3].

We are not going to give a formal definition of $\mathcal{R}$-spaces and refer the reader to [PK] for details. Roughly speaking, a linear $\mathcal{R}$-space is a linear space X of functions on $\mathbb{T}$ such that $X \subset VMO$ and there exists a Köthe sequence space E such that $\varphi \in X$ if and only if $\varphi \in BMO$ and the singular values of the Hankel operators H_φ and $H_{\bar{\varphi}}$ satisfy

$$\{s_n(H_\varphi)\}_{n\ge 0} \in E \quad \text{and} \quad \{s_n(H_{\bar{\varphi}})\}_{n\ge 0} \in E.$$

Recall that E is a *Köthe sequence space* if

$$\{c_n\}_{n\ge 0} \in E, \quad |d_n| \le |c_n| \quad \Rightarrow \quad \{d_n\}_{n\ge 0} \in E.$$

Important examples of $\mathcal{R}$-spaces are the Besov spaces $B_p^{1/p}$, $0 < p < \infty$ (see [PK], [Pe1] (the corresponding Köthe space E is ℓ^p) and the space VMO of functions of vanishing mean oscillation (E is the space c_0 of sequences converging to 0). We refer the reader to [G] for definitions and properties of the spaces BMO and VMO.

The second class consists of function spaces $X \subset C(\mathbb{T})$ that contain the trigonometric polynomials and satisfy the following axioms:

(A1) *If $f \in X$, then $\bar{f} \in X$ and $\mathbb{P}_+ f \in X$;*

(A2) *X is a Banach algebra with respect to pointwise multiplication;*

(A3) *for every $\varphi \in X$ the Hankel operator H_φ is a compact operator from X_+ to X_-;*

(A4) *if $f \in X$ and f does not vanish on $\mathbb{T}$, then $1/f \in X$.*

Here we use the notation

$$X_+ = \{f \in X :\ \hat{f}(j) = 0,\ j < 0\}, \quad X_- = \{f \in X :\ \hat{f}(j) = 0,\ j \ge 0\}.$$

For simplicity we write $\Phi \in X$ if all entries of a matrix function Φ belong to X.

Note here that the Besov classes $B^s_{p,q}$, $1 \le p < \infty$, $1 \le q \le \infty$, $s > 1/p$, the space of functions with absolutely convergent Fourier series, the spaces

$$\{f :\ f^{(n)} \in VMO\}, \quad \{f :\ \mathbb{P}_+ f^{(n)} \in C(\mathbb{T}),\ \mathbb{P}_- f^{(n)} \in C(\mathbb{T})\}, \quad n \ge 1,$$

satisfy (A1)–(A4).

Among nonseparable Banach spaces X satisfying (A1)–(A4) we mention the following ones: the Hölder–Zygmund spaces Λ_α, $\alpha > 0$, the spaces

$$\{f :\ f^{(n)} \in BMO\}, \quad \{f :\ \mathbb{P}_+ f^{(n)} \in H^\infty,\ \mathbb{P}_- f^{(n)} \in H^\infty\}, \quad n \ge 1,$$

the space

$$\{f :\ |\hat{f}(j)| \le \text{const}(1 + |j|)^{-\alpha}\}, \quad \alpha > 1,$$

(see [Pe3]).

We need the following so-called recovery theorem for unitary-valued functions. Let X be either an $\mathcal{R}$-space or a space of functions satisfying (A1)–(A4) and let

$\mathcal{U}$ be a unitary-valued matrix function in $X(\mathbb{M}_{n,n})$ such that the Toeplitz operator $T_{\mathcal{U}} : H^2(\mathbb{C}^n) \to H^2(\mathbb{C}^n)$ has dense range. Then

$$\mathbb{P}_{-}\mathcal{U} \in X \quad \Longrightarrow \quad \mathcal{U} \in X. \tag{9.1}$$

Moreover, if X is a Banach $\mathcal{R}$-space, then we can conclude that

$$\|\mathbb{P}_{-}\mathcal{U}\|_X \le \text{const}\,\|\mathcal{U}\|_{\mathcal{X}}.$$

For $\mathcal{R}$-spaces this was proved in [Pe2], for spaces satisfying (A1)–(A4) this was proved in [Pe3], see also [PK] for the scalar case. In fact, in [PK] and [Pe3] the assumptions on X were slightly different but it was shown in [AP2] that the method used in [Pe3] can be adjusted to work for all spaces satisfying (A1)–(A4).

Note that both $\mathcal{R}$-spaces and spaces satisfying (A1)–(A4) are contained in VMO. Therefore if $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and $\mathbb{P}_{-}\Phi \in X(\mathbb{M}_{m,n})$, then H_Φ is compact.

To establish the main result of this section we need the fact that for a linear $\mathcal{R}$-space X the space $X \cap L^\infty$ forms an algebra with respect to pointwise multiplication. Indeed, suppose that $\varphi, \psi \in X \cap L^\infty$. Clearly, $\varphi \in X$ if and only if $\bar{\varphi} \in X$, and so it is sufficient to prove that $\mathbb{P}_{-}(\varphi\psi) \in X$. Let $f \in H^2$. We have

$$H_{\varphi\psi} f = \mathbb{P}_{-}\varphi\psi f = \mathbb{P}_{-}\varphi\mathbb{P}_{+}\psi f + \mathbb{P}_{-}\varphi\mathbb{P}_{-}\psi f = H_\varphi T_\psi f + \breve{T}_\varphi H_\psi f, \tag{9.2}$$

where the operator $\breve{T}_\varphi : H^2_{-} \to H^2_{-}$ is defined by $\breve{T}_\varphi g = \mathbb{P}_{-}\varphi g$, $g \in H^2_{-}$. It is easy to deduce from (9.2) that $\mathbb{P}_{-}(\varphi\psi) \in X$.

Theorem 9.1. *Suppose that X is a linear $\mathcal{R}$-space or X is a function space satisfying* (A1)–(A4). *Let Φ be a bounded $m \times n$ matrix function such that $\mathbb{P}_{-}\Phi$ is a nonzero matrix function in $X(\mathbb{M}_{m,n})$. If $F \in H^\infty(\mathbb{M}_{m,n})$ is a best approximation of Φ and $\Phi - F$ admits a partial canonical factorization*

$$\Phi - F = \begin{pmatrix} \overline{\Omega} & \Xi \end{pmatrix} \begin{pmatrix} t_0 U & 0 \\ 0 & \Psi \end{pmatrix} \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}^*,$$

then $\Upsilon, \Theta, \Omega, \Xi, U, \mathbb{P}_{-}\Psi \in X$.

Proof. Assume without loss of generality that $t_0 = 1$. By Theorem 3.4, if we replace a best approximating function F with any other best approximation, we do not change $\mathbb{P}_{-}\Psi$. Thus we may assume that F is the unique superoptimal approximation of Φ by bounded analytic matrix functions. Then $\Phi - F$ belongs to X. For linear $\mathcal{R}$-spaces this was proved in [PY1], Theorem 5.1. For spaces X satisfying (A1)–(A4) this is Theorem 9.2 of [Pe3]. (Theorem 9.2 is stated in [Pe3] under slightly different assumptions but using the results of §4 of [AP2], one can easily see that the proof given in [Pe3] works for all spaces satisfying (A1)–(A4).)

Let us first prove that $\Theta \in X$. Consider a partial thematic factorization of $\Phi - F$ of the form (5.2). Then the matrix functions V_j given by (5.3) belong to X. If X is a linear $\mathcal{R}$-space, this was proved in §5 of [PY1], for spaces satisfying (A1)–(A4) this was proved in §9 of [Pe3]. (Again this was proved in [Pe3] under

slightly different assumptions but the results of §4 of [AP2] show that the proof given in [Pe3] works for all spaces satisfying (A1)–(A4).) In particular the inner matrix functions Θ_j in (5.3) belong to X. Since $X \cap L^\infty$ is an algebra, it follows now from (5.4) that $\Theta \in X$.

Consider now the unitary-valued matrix function $\mathcal{V} = \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix}$. We have just proved that $\mathbb{P}_-\mathcal{V} \in X$. As we have mentioned in §3, the Toeplitz operator $T_{\mathcal{V}}$ has dense range in $H^2(\mathbb{C}^n)$. Therefore by (9.1), $\mathcal{V} \in X$, and so $\Upsilon \in X$.

If we apply the above reasoning to $\Phi^{\rm t}$, we prove that $\Xi \in X$ and $\Omega \in X$. It follows now from (4.7) that $U \in X$. Finally, it follows from (5.8) that $\mathbb{P}_-\Psi \in X$. ■

Remark. It can be shown that if X is a linear $\mathcal{R}$-space, then the X-norms of $\Upsilon, \Theta, \Omega, \Xi, U, \mathbb{P}_-\Psi$ can be estimated in terms of the X-norm of $\mathbb{P}_-\Phi$.

Clearly, it follows from Theorem 9.1 that the same result holds for arbitrary partial canonical factorizations. In particular, the following theorem holds.

Theorem 9.2. *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and F is the superoptimal approximation of Φ by bounded analytic matrix functions. If* (7.5) *is a canonical factorization of $\Phi - F$, then all factors on the right-hand side of* (7.5) *belong to X.*

Consider now separately the important case $X = VMO$. As we have already noted, VMO is an $\mathcal{R}$-space. It is well known that if $\Phi \in L^\infty$, then $\mathbb{P}_-\Phi \in VMO$ if and only if $\Phi \in H^\infty + C$. It is also well known that $QC = VMO \cap L^\infty$.

Theorem 9.3. *Let $\Phi \in (H^\infty + C)(\mathbb{M}_{m,n})$ and $\mathbb{P}_-\Phi \neq 0$. If F is a best approximation of Φ by bounded analytic functions and $\Phi - F$ admits a partial canonical factorization* (4.6)*, then $\mathcal{V}, \mathcal{W}, U \in QC$ and $\Psi \in H^\infty + C$.*

Theorem 9.3 follows immediately from Theorem 9.1 if we put $X = VMO$.

10. Very badly approximable unitary-valued functions

In this section we study unitary-valued very badly approximable matrix functions which play a crucial role in canonical factorizations. The following theorem describes such functions U under the condition $\|H_U\|_{\rm e} < 1$.

Theorem 10.1. *Let U be an $r \times r$ unitary-valued matrix function such that $\|H_U\|_{\rm e} < 1$. The following are equivalent:*
(i) *U is very badly approximable;*
(ii) *the Toeplitz operator $T_{zU} : H^2(\mathbb{C}^r) \to H^2(\mathbb{C}^r)$ has dense range in $H^2(\mathbb{C}^r)$;*
(iii) *the Toeplitz operator $T_{\bar{z}U^*} : H^2(\mathbb{C}^r) \to H^2(\mathbb{C}^r)$ has trivial kernel.*

Proof. First of all it is trivial that (ii) is equivalent to (iii). The implication (i)⇒(ii) is proved in the Remark after the proof of Theorem 4.3.

It remains to show that (ii) implies (i). Again, it is explained in the Remark following the proof of Theorem 4.3 that T_U is Fredholm. Consider a Wiener–Hopf factorization of U:

$$U = \Psi_2^* \begin{pmatrix} z^{d_1} & 0 & \cdots & 0 \\ 0 & z^{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z^{d_r} \end{pmatrix} \Psi_1,$$

where $\Psi_1^{\pm 1}, \Psi_2^{\pm 1} \in H^2(\mathbb{M}_{r,r})$. It is well known and it is easy to see that (ii) is equivalent to the condition that all Wiener–Hopf indices d_j are negative. It follows from the results of §3 of [AP2] that the superoptimal singular values $t_0(U), \cdots, t_{r-1}(U)$ are equal to 1 which means that U is very badly approximable. ■

References

[AAK1] V.M. Adamyan, D.Z. Arov, and M.G. Krein, On infinite Hankel matrices and generalized problems of Carathéodory-Fejér and F. Riesz, *Funktsional. Anal. i Prilozhen.* **2:1** (1968), 1-19 (In Russian).

[AAK2] V.M. Adamyan, D.Z. Arov, and M.G. Krein, On infinite Hankel matrices and generalized problems of Carathéodory-Fejér and I. Schur, *Funktsional. Anal. i Prilozhen.* **2:2** (1968), 1-17 (In Russian).

[AAK3] V. M. Adamyan, D. Z. Arov, and M. G. Krein, Infinite Hankel block matrices and some related continuation problems, *Izv. Akad. Nauk Armyan. SSR, Ser. Mat.,* **6** (1971), 87-112 (Russian)

[AP1] R.B. Alexeev and V.V. Peller, Invariance properties of thematic factorizations of matrix functions, to appear

[AP2] R.B. Alexeev and V.V. Peller, Unitary interpolants and factorization indices of matrix functions, to appear

[D] R.G. Douglas, "Banach algebra techniques in operator theory", Academic Press, New York–London 1972.

[F] B. A. Francis, "A Course in H^∞ Control Theory", Lecture Notes in Control and Information Sciences No. 88, Springer Verlag, Berlin, 1986.

[G] J. Garnett, "Bounded analytic functions", Academic Press, NY–London–Toronto–Sydney–San Franciso, 1981.

[Kh] S. Khavinson, On some extremal problems of the theory of analytic functions, *Uchen. Zapiski Mosk. Universiteta, Matem.* **144**:4 (1951), 133-143. English Translation: Amer. Math. Soc. Translations (2) **32** (1963), 139-154.

[LS] G.S. Litvinchuk and I.M. Spitkovski, "Factorization of measurable matrix functions", Oper. Theory: Advances and Appl., **25**. Birkhäuser Verlag, Basel-Boston, MA, 1987.

[Ne] Z. Nehari, On bounded bilinear forms, *Ann. Math.,* **65** (1957), 153-162.

[Ni] N.K. Nikol'skii, "Treatise on the shift operator. Spectral function theory," Springer-Verlag, Berlin–Heidelberg–New York–Tokyo, 1986.

[Pa] L. B. Page, Bounded and compact vectorial Hankel operators, *Trans. Amer. Math. Soc.*, **150** (1970) 529-539.

[Pe1] V.V. Peller, A description of Hankel operators of class $\mathfrak{S}_p$ for $p > 0$, investigation of the rate of rational approximation and other applications, *Mat. Sb.* **122** (1983), 481-510; English Translation: *Math. USSR-Sb.* **50** (1985), 465-494.

[Pe2] V.V. Peller, Hankel operators and multivariate stationary processes, *Proc. Symp. Pure Math.* **51** (1990), 357-371.

[Pe3] V.V. Peller, Factorization and approximation problems for matrix functions, *Amer. J. Math.* **11** (1998), 751-770.

[Pe4] V.V. Peller, Hereditary properties of solutions of the four block problem, Indiana Univ. Math. J. **47** (1998), 177-197.

[PK] V.V. Peller and S.V. Khruschev, Hankel operators, best approximation and stationary Gaussian processes, *Uspekhi Mat. Nauk* **37:1** (1982), 53-124. English Translation: *Russian Math. Surveys* **37** (1982), 53–124.

[PT] V.V. Peller and S.R. Treil, Approximation by analytic matrix functions. The four block problem, *J. Funct. Anal.* **148** (1997), 191-228.

[PY1] V.V. Peller and N.J. Young, Superoptimal analytic approximations of matrix functions, *J. Funct. Anal.* **120** (1994), 300-343.

[PY2] V.V. Peller and N.J. Young, Superoptimal singular values and indices of matrix functions, *Int. Eq. Op. Theory* **20** (1994), 35-363.

[PY3] V.V. Peller and N.J. Young, Continuity properties of best analytic approximation, *J. Reine Angew. Math.* **483** (1997), 1-22.

[Po] S. J. Poreda, A characterization of badly approximable functions, *Trans. Amer. Math. Soc.*, **169** (1972), 249-256.

[Sa] D. Sarason, "Function theory on the unit circle," Notes for lectures at Virginia Polytechnic Inst. and State Univ., 1978.

[Si] I.B. Simonenko, Some general problems of the theory of the Riemann boundary-value problem, *Izv. Akad. Nauk SSSR, Ser. Mat.* **32** (1968), 1138-1146 (In Russian).

[SNF1] B. Sz.-Nagy and C. Foias, "Harmonic analysis of operators on Hilbert space", Akadémiai Kiadó, Budapest, 1970.

[SNF2] B. Sz.-Nagy and C. Foias, On contractions similar to isometries and Toeplitz operators, *Ann. Acad. Sci.. Fenn.,* A I **2** (1976), 553-564.

[T] S.R. Treil, On superoptimal approximation by analytic and meromorphic matrix-valued functions, *J. Functional Analysis* **131** (1995), 386-414.

[V] V.I. Vasyunin, Formula for multiplicity of contractions with finite defect indices, *Oper. Theory: Adv. Appl.*, Birkhäuser **4** (1989), 281-304.

[Y] N. J. Young, The Nevanlinna-Pick problem for matrix-valued functions, *J. Operator Theory* **15** (1986), 239-265.

Department of Mathematics
Kansas State University
Manhattan, Kansas 66506
USA